Variational segmentation problems using prior knowledge in imaging and vision

Fundana, Ketut

Publication date:
2010

Document version
Tidlig version også kaldet pre-print

Citation for published version (APA):
Fundana, K. (2010). *Variational segmentation problems using prior knowledge in imaging and vision*. Department of Computer Science, Faculty of Science, University of Copenhagen.

Variational Segmentation Problems using Prior Knowledge in Imaging and Vision

Dissertation

Ketut FUNDANA

Image Group
Department of Computer Science
Faculty of Science

UNIVERSITY OF COPENHAGEN

Supervisors
Prof. Mads Nielsen
Prof. Anders Heyden

January 2010

This dissertation is a result of cooperation between the Applied Mathematics Group of School of Technology at Malmö University, Sweden, and the Image Group of Department of Computer Science at University of Copenhagen, Denmark.

The research was funded by the COMPUTATIONAL AND COGNITIVE VISION SYSTEMS: VISIONTRAIN RTN-CT-2004-005439 Marie Curie Action within the EC's Sixth Framework Programme.

Abstract

This dissertation addresses variational formulation of segmentation problems using prior knowledge. Variational models are among the most successful approaches for solving many Computer Vision and Image Processing problems. The models aim at finding the solution to a given energy functional defined to describe a Computer Vision task through energy minimization.

Image segmentation, as an ill-posed problem, is still a major challenge in Computer Vision. Due to the presence of noise, clutter and occlusion, the use of image information alone often gives poor segmentation results. To overcome this problem, prior knowledge is needed to obtain the desired solution. The introduction of shape priors in particular, has proven to be an effective way to segment objects of interests.

Firstly, we propose a prior-based variational segmentation model to segment objects of interest in image sequences, that can deal with shape deformations and at the same time is robust to noise, clutter and occlusions. The proposed model is based on the Chan-Vese functional coupled with a frame-to-frame interaction term as a shape prior. In order to deal with severe occlusions, we combine the prior-based segmentation model with a novel variational contour matching algorithm in order to detect and locate the occlusion. By having information about the occlusion, the segmentation results can be improved.

Variational segmentation models suffer from the existence of local minima due to the non-convexity of the energy functionals. The non-convex functionals are usually minimized by methods based on gradient descent, which leads to undesired solutions. This makes the initialization critically important to get satisfactory results. To overcome the local minima problem, we propose to use a variational segmentation model in a global minimization framework by convexifying the Chan-Vese model to obtain the global minimizers. This convex formulation can be regarded

as a continuous counterpart of the Graph-cuts in the discrete segmentation models. The convex model is then extended by adding a shape prior term in order to segment the object of interest.

Many objects have high variability in shape and orientation. This often leads to unsatisfactory results, when using a segmentation model with single shape template. One way to solve this is by using more sophisticated shape models. We propose to incorporate shape priors from a shape sub-manifold of pose-invariant planar contours into both the Chan-Vese model and its convex formulation to segment an object of interest in a sequence of images. We apply the models to track the viewpoint onto 3D rigid object.

The prior-based object segmentation models encounter the problem of shape alignment, where pose invariant parameters complicate the optimization of the model. To overcome the common numerical problems associated with the step size of the pose parameters in the discretization of the pose model, we propose a novel gradient procedure for the pose estimation based on the construction of the Riemannian structure on the group of transformations and a derivation of the corresponding pose energy gradient.

Finally, we show that the convex energy functional can be extended and used for segmenting data on manifolds into multi-regions with constant properties. We implement the adaptation of the model to data that is represented by triangular meshes.

Preface

This dissertation is submitted to the University of Copenhagen as a partial fulfilment of the requirements for the degree of Doctor of Philosophy (PhD) in Computer Science. This dissertation is written based on the following work:

- Amaäl Delaunoy and Ketut Fundana and Emmanuel Prados and Anders Heyden. Convex Multi-region Segmentation on Manifolds. In *Proceeding of International Conference on Computer Vision (ICCV)*, Kyoto, 2009 [48].
- Niels Chr. Overgaard and Ketut Fundana and Anders Heyden. Pose Invariant Shape Prior Segmentation using Continuous Cuts and Gradient Descent on Lie Groups. In *Proceeding of International Conference on Scale Space and Variational Methods (SSVM)*, Voss, 2009 [112].
- Ketut Fundana and Anders Heyden and Christian Gosch and Christoph Schnoerr. Continuous graph cuts for prior-based object segmentation. In *Proceeding of International Conference on Pattern Recognition (ICPR)*, Tampa, 2008 [64].
- Christian Gosch and Ketut Fundana and Anders Heyden and Christoph Schnoerr. View Point Tracking of Rigid Objects Based on Shape Sub-manifolds. In *Proceeding of European Conference on Computer Vision (ECCV)*, Marseille, 2008 [73].
- Ketut Fundana and Niels Chr. Overgaard and Anders Heyden. Variational Segmentation of Image Sequences Using Region-Based Active Contours and Deformable Shape Priors. *International Journal of Computer Vision, 80(3), pp. 289-299*, 2008 [65].
- Ketut Fundana and Niels Chr. Overgaard and Anders Heyden and David Gustavsson and Mads Nielsen. Nonrigid Object Segmentation and Occlusion Detection in Image Sequences. In *Proceeding of International Conference on Computer Vision Theory and Application (VISAPP)*, Madeira, 2008 [68].

- David Gustavsson and Ketut Fundana and Niels Ch. Overgaard and Anders Heyden and Mads Nielsen. Variational Segmentation and Contour Matching of Non-Rigid Moving Object. In *Proceeding of Internatinal Conference on Computer Vision (ICCV) Workshop on Dynamical Vision*, Rio de Janeiro, 2007 [74].
- Ketut Fundana and Niels Ch. Overgaard and Anders Heyden. Deformable Shape Priors in Chan-Vese Segmentation of Image Sequences. In *Proceeding of International Conference on Image Processing*, San Antonio, 2007 [66].
- Ketut Fundana and Niels Ch. Overgaard and Anders Heyden. Variational Segmentation of Image Sequences using Deformable Shape Prior. In *Proceeding of Scandinavian Conference on Image Analysis*, Aalborg, 2007 [67].

Acknowledgements

During my PhD studies, I had an opportunity to work closely with a number of people. First of all, I would like to express my gratitude to my supervisors, Mads Nielsen and Anders Heyden for their guidance and support. I would like to thank Niels Chr. Overgaard for the collaborations and his contributions to this dissertation. I would also like to thank those people who also contributed to this dissertation: David Gustavsson for the collaboration during his visit at Applied Mathematics Group at Malmö University, Christoph Schnoerr and Christian Gosch for the excellent collaboration and discussions during my visit at IPA/CVGPR Group of University of Heidelberg/Mannheim, Emannuel Prados and Amaël Delaunoy for our interesting work and discussions during my visit at PERCEPTION Group of INRIA Rhône-Alpes in Grenoble. I would like to thank Radu Horaud as the project coordinator and other VISIONTRAIN project members. I would also like to thank Mattias Hansson, Alma Masic, Melanie Ganz and Aasa Feragen for proofreading and helpful comments, and other former and current colleagues at Applied Mathematics Group of Malmö University and the Image Group of University of Copenhagen. Finally I would like to thank my family and Sofia for their love and support.

Contents

List of Figures **xv**

1 Introduction **1**

1.1 Background 1

1.2 Segmentation Models 3

1.2.1 The Active Contour Models 5

1.2.2 Convex Formulation of the Active Contour Models 10

1.3 Prior Knowledge in Segmentation Problems 13

1.3.1 Total Variation 14

1.3.2 Shape Priors 15

1.4 Contributions/Summaries of the Papers 17

2 Variational Segmentation of Image Sequences using Region-Based Active Contours and Deformable Shape Priors **25**

2.1 Introduction 25

2.2 Theoretical Background 27

2.2.1 Region-Based Segmentation 27

2.2.2 The Level Set Method 28

2.2.3 Gradient Descent Evolution 29

2.3 Remarks about Minimization of the Chan-Vese Functional 30

2.4 Segmentation of Image Sequences 31

2.4.1 A Variational Updating-Model 31

2.4.2 The Interaction Term 32

2.4.3 The Gradient Descent Equations 33

2.5 Remarks about the Interaction Term 34

2.5.1 Area of the Symmetric Difference . . . 35
2.5.2 The Pseudo-Distance . . . 36
2.5.3 Proposed Interaction Term . . . 37
2.6 Experimental Results . . . 37
2.7 Conclusions . . . 39

3 Nonrigid Variational Object Segmentation with Occlusion Detection in Image Sequences 43
3.1 Introduction . . . 44
3.2 Segmentation of Image Sequences . . . 45
3.2.1 Region-Based Segmentation . . . 45
3.2.2 The Interaction Term . . . 46
3.2.3 Using the Interaction Term in Segmentation of Image Sequences . . . 47
3.3 Occlusion Detection by Contour Matching . . . 48
3.3.1 A Contour Matching Problem . . . 49
3.3.2 Occlusion Detection . . . 52
3.3.3 Re-segmentation . . . 52
3.4 Experimental Results . . . 53
3.5 Conclusions . . . 56

4 Continuous Cuts for Prior-Based Object Segmentation 59
4.1 Introduction . . . 59
4.2 Continuous Cuts . . . 60
4.3 Shape Priors for Continuous Cuts . . . 61
4.4 Implementation and Results . . . 64
4.5 Conclusions . . . 65

5 View Point Tracking of Rigid Objects using Region-Based Segmentation Model and Shape Priors from Shape Submanifolds 67
5.1 Introduction . . . 67
5.2 Shape Model, Object Representation, Learning . . . 69
5.3 Pose Inference and Tracking on the View Sphere . . . 70
5.4 Segmentation and Image Contours . . . 74
5.5 Experiments and Evaluation . . . 76

5.6 Conclusions and Further Work 76

6 Pose Invariant Shape Prior Segmentation using Continuous Cuts and Gradient Descent on Lie Groups 81

6.1 Introduction 82
6.2 Background: Relaxation in the CV Model 84
6.2.1 Relaxation in the CV Model 84
6.2.2 Continuous Cuts for Prior-Based Object Segmentation 86
6.3 The Shape Prior Segmentation Model 86
6.3.1 The Basic Energy Functional 86
6.3.2 Pose Invariant Prior Interaction Energy 88
6.3.3 The Gradient Construction 89
6.4 Experiments 91
6.5 Conclusions 92

7 Convex Multi-Region Segmentation on Manifolds 95

7.1 Introduction 95
7.1.1 Global Multi-Region Segmentation 96
7.1.2 Data Segmentation on Manifolds 97
7.1.3 Contributions 98
7.2 Multi-Region Segmentation Model 98
7.2.1 Convex Two-Phases Model 98
7.2.2 Extension to Multi-Region Segmentation 99
7.3 Multi-Region Segmentation on Manifolds 99
7.3.1 Optimization Method 101
7.3.2 Applications 102
7.4 Experiments 104
7.4.1 The Two Region Case 104
7.4.2 Dealing with Multiple Regions 107
7.5 Conclusions 110

8 Discussion and Future Work 113

References 117

List of Figures

1.1 Different Type of Images. 2

1.2 The curve Γ is represented by the zero-level set of the function ϕ. 4

1.3 An example of contour evolution in image segmentation using the Chan-Vese model. 9

1.4 Two level set functions are needed to segment four regions using the multi-phase model of Vese and Chan [147]. 9

1.5 Segmentation of a horse with the Chan-Vese model (left) and convex formulation of the Chan-Vese model (right). 12

1.6 Segmentation of a horse in Figure 1.5 is improved by adding a shape prior term. 17

2.1 The initialization (green contour) and segmentation results of a person (red contour) in an image sequence using the Chan-Vese model. 32

2.2 The initialization (green contour) and segmentation results (red contour) of a nonrigid object in a synthetic image sequence with additive Gaussian noise. First, second, and third rows: without interaction term and increasing α. Fourth row: with interaction term. 38

2.3 The initialization (green contour) and segmentation results (red contour) of a car which passes through a lamp post in a traffic sequence. First and second columns: without interaction term and increasing α. Third and fourth columns: with interaction term and increasing α. 40

2.4 Segmentation results taken from the fourth row of Figure 2.3. Left: without interaction term. Right: with interaction term. 41

2.5 The initialization (green contour) and segmentation results (red contour) of a person that partly occluded by another person in a human walking sequence. First and second columns: without interaction term and increasing α. Third and fourth columns: with interaction term and increasing γ. 42

3.1 Given two closed curves Γ_1 and Γ_2 contained in two images F_1 and F_2, Φ maps F_1 onto F_2 such that Γ_1 is mapped onto Γ_2. 49
3.2 Segmentation of a nonrigid object in a synthetic image sequence with additive Gaussian noise (Frame 1-6). Top Row: without the interaction term, noise in the occlusion is captured. Bottom Row: with interaction term, we obtain better results. 54
3.3 Left: Deformation field. Right: Frame 3 after deformation according to the displacement field onto Frame 4. 54
3.4 Detected occlusions in Frame 2-5 of Figure 3.2. 55
3.5 Segmentation of a walking person partly covered by an occlusion in the human walking sequence (Frame 1-5). Top Row: without interaction term, and Bottom Row: with interaction term 55
3.6 Detected occlusion in Frame 2-3 of Figure 3.5. 55
3.7 Segmentation of the synthetic image sequence by using smaller coupling constant than the one in Figure 3.2. Top row: without reconstruction of the occluded regions. Bottom row: after the occluded regions are reconstructed. . . . 56
3.8 Segmentation of the human walking sequence when by using smaller coupling constant than the one in Figure 3.5. Top row: without reconstruction of the occluded regions. Bottom row: after the occluded region is reconstructed. . . . 56

4.1 Motivation for using the shape prior as in (4.4) 62
4.2 The evolution of $u(x)$ (top row) and the corresponding histogram (bottom row). First column: initial. Middle columns: intermediate results. Right column: final results 63
4.3 Segmentation of a bird using continuous graph cuts and its corresponding u. Top row: without a shape prior. Bottom row: with a shape prior. 66
4.4 Segmentation of a cup using continuous graph cuts and its corresponding u. Top row: without a shape prior. Bottom row: with a shape prior. 66

5.1 Illustration of a view sphere. Right hand: indicated are three sampled contours of an airplane seen from a camera from points on the view sphere. The object is located in the center of the sphere. Left hand: illustration of the shape submanifold. The green lines between sphere and manifold indicate corresponding points, the blue arrow indicates a point that is interpolated using, in this case, three points which are neighbors on the sphere. This specific object was taken from the Princeton 3D shape benchmark [135]. 70

5.2 Representing and tracking shape changes as motions on the view sphere. Blue: measurements P_k. Red: path $s(t)$ of the mass point. Magenta: predicted points. The start point of the trajectory is at the far left end. The green grid lines indicate the underlying sphere. 71

5.3 Illustration of shape interpolation with Karcher means in the closed pre-shape space $\mathcal{C}$. The corners represent the original shapes, the other contours are interpolations weighted with their barycentric coordinates. The corner curves are randomly chosen from the MPEG-7-CE1 shape data base. 72

5.4 Keeping track of the spherical position: Shape c_k and position t_k are known, as well as a new shape q. What is the (approximate) position t_{k+1} on the view sphere corresponding to q? 72

5.5 Experiment tracking the view sphere position using only the segmented contours from a sequence of images. Right: shown are measurements obtained on the view sphere, for the complete sequence. Left: a few images from the sequence are shown, the corresponding interpolated contours from the shape space $\mathcal{C}$ to their right. The initial position $t_0 \in \mathbb{S}^2$ and shape s_0 were given manually. Then for each image, the result from the previous one was used as initialization. A region based level set segmentation was used, with a curvature regularization term after [51]. 77

5.6 Sphere tracking experiment with occlusion. The top row shows the tracked view sphere path on the right (the arrows indicate the direction of motion), and an illustration of the image sequence on the left. The color coding shows the corresponding contours and view sphere positions. Using the resulting shape from each previous frame to create a prior for the segmentation algorithm enables the sphere tracking to keep going for this sequence, where a small occluding object moves in front of the object. Each row shows the area of interest from 3 subsequent frames with the superimposed segmentation result, followed by the contour representing the shape tracked on the view sphere. 78

5.7 Sphere tracking with a real recorded sequence totalling 97 frames. Roughly every 20^{th} is shown. Indicated in each frame are the segmentation result (green) and aligned interpolated shape (red). Difficult situations where the view tracking goes wrong are indicated in red, yellow are situations which are just ok. The time line on the bottom indicates the situation for the whole 97 frames.The spheres on the right indicate the inferred view positions along the sequence. . . 79

6.1 Experiment 1: First row: The original image, 212×320 pixels (left), the active contour $\Gamma = \{x; u(x) = .5\}$ in CV-segmentation without priors after 100 iterations (middle), and the corresponding segmentation (right). Second row: The shape template, the active contour and the shape prior after 150 iterations, and the final segmentation. Final row: segmentation of the image contaminated with 15% Gaussian noise using 200 iterations. Parameters: $\mu = .4$, $\lambda = .1$, $\theta = .5$ and step-size $\Delta t = .75$. 92

6.2 Experiment 2: Shape prior segmentation with three different initial poses (top row). Evolution after (approximately) 12, 25, 50, 100 and 200 iterations (rows 2–6). The run-time for 100 iterations is about 25 CPU-seconds. In the final phase of the segmentation, object previously detected outside the prior disappears. With the third initialization the he shape prior gets stuck in a local minimum. Such behavior cannot be ruled out when we work with local optimization methods. Image size and parameter settings are as in Experiment 1. 94

7.1 Example of segmentation on a manifold. (a) The Input textured surface. (b) The triangular representation of the surface with the retrieved contour (in red). (c) Surface colored with the mean values of the segmented regions (and surface shading). 96
7.2 Segmentation result on the synthesized Stanford bunny surface. (a) Input shaded object. (b) Input mesh with synthetic texture mapping. (c) Input textured mesh (shaded visualization) and final contour (in red). 105
7.3 The evolution on the synthesized bunny surface. Different initialization of **U** (first row); Intermediate values of **U** (second row); The obtained solution **U** (third row); The obtained mean values (fourth row) with shading. 106
7.4 Segmentation results on meshes in the two-region segmentation on three different examples. From top to bottom: Input textured mesh; Mesh shape where the segmentation is performed and the initial random value of one component of **U**; Recovered mean values of each region; Segmented object. 107
7.5 Segmentation results on meshes in the multi-region case on synthesized examples inspired from [91]. (a) Input textured mesh (same shape as previous Butterfly and Horses data). (b) Clustering using K-Means algorithm. (c) Recovered mean values of each region obtained by our approach. 108
7.6 Segmentation results on meshes in the multi-region case. Top row: Horse data set and its segmentation for three regions. Bottom row: Four regions labeling of the Butterfly data. (a) Input textured mesh; (b) Recovered mean values of each region obtained by our approach; (c) One of the segmented regions. . . . 109
7.7 Segmentation results on mesh curvature in 3 regions. (a) Input mesh. (b) Mean Curvature visualization. (c) Simple thresholding of the mean curvature. (d) Segmentation result of the mean curvature into three regions with our approach. 110
7.8 Segmentation result on a colored mesh obtained by multiview stereo algorithm. Front view (top row) and back view (bottom row). Original input colored mesh and the associated 3D shape (left). Result of the segmentation into three regions obtained by our algorithm (middle). Recovered mean values displayed for each region (right). 111

GLOSSARY

TV	Total Variation
PDE	Partial Differential Equation
ODE	Ordinary Differential Equation
PCA	Principal Component Analysis
CV	Chan and Vese Segmentation Model
VC	Vese and Chan Multiphase Segmentation Model
ROF	Rudin, Osher and Fatemi Model for Denoising
GAC	Geodesic Active Contour Segmentation Model
MS	Mumford-Shah Model

Chapter 1

Introduction

1.1 Background

Modern life has been greatly influenced by computer vision and image processing from digital communication, entertainment, medicine, security, and so on. The fast development and wide use of imaging science and technology such as video cameras, satellite imaging, Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Radar, Ultrasound, and more, have provided huge amount of data. As a consequence, the need for challenging applications/processes in artificial vision has increased which demands more efficient, accurate and stable mathematical models and algorithms to process and interpret the data. The range of applications of computer vision and image processing is large including tracking of objects in image sequences, scene understanding, image retrieval, object recognition, and segmentation to name a few. Those applications or processes involved can be classified into three categories: low-level vision, which includes basic operation such as edge detection, mid-level vision, such as segregating objects from the background and high-level vision that gives semantic meaning to the images such as face recognition.

The extraction of information from images is highly difficult due to the complexity and diversity of images, even for human vision, for example: finding relationship between 3D world and 2D image plane, inferring the depth, finding how pixels are grouped and segmented, and etc. Mathematically speaking, the problems in computer vision and image processing are ill-posed inverse problems.

In this dissertation, we are dealing with the segmentation problem which is one of the challenging processes in computer vision and image processing. This important process provides crucial information for a variety of high level applications, such as 3D reconstruction, video surveillance, object recognition, activity recognition, etc. The objective of segmentation is to partition a given image into semantically meaningful regions or objects with consistent properties. This is a very difficult problem to solve with many challenges depending on the data and applications especially since the definition of objects or regions is subjective. In the case of object segmentation, for instance, although the object of interest in a given image is known, it may change its pose, viewpoint, illumination, etc., which puts additional constraints on the segmentation process. Moreover, the presence of noise, clutter and occlusions, which commonly appear, can lead to inconsistent segmentation results.

There are several cues which can be used to distinguish an object from the background in a given image. Low-level information or local image features (edges, gray values, etc) that follow some basic assumptions, such as gray level uniformity or edge coherence, can be used for segmentation. However, in many cases, those assumptions are usually insufficient to segment objects of interest, for example due to missing or misleading information. Therefore, higher-level information such as motion, color distribution, smoothness, texture, pose, shape, etc., are needed to make the segmentation models more robust. As in human vision which tends to integrate low-level and high-level information, most real-life applications need to have some prior knowledge about the regions or objects of interest beforehand. Figure 1.1 shows some examples of images for segmentation that require some prior knowledge. Consider the ultrasound example (d). Segmentation of the heart chamber is not an easy task since the edges are not obvious and the intensities can mislead the result. Therefore, prior knowledge such as motion, shape and information about the patient (age, gender, health conditions, etc.) are needed.

(a) Natural **(b)** Texture **(c)** MRI **(d)** Ultrasound

Figure 1.1: Different Type of Images.

Due to the importance of the prior knowledge in the segmentation process, mathematical modeling of prior knowledge and its integration into segmentation models become critical. We are in particular interested in segmentation of meaningful regions or objects using some prior knowledge such as region characteristics, occlusion information, smoothness in term of the total variation and the shape of the objects.

1.2 Segmentation Models

Many approaches have been proposed to solve the segmentation problems such as thresholding, clustering, region-growing, split and merge, watershed, statistical models, variational methods, and graph-based approaches to name a few. The survey of all approaches is beyond the scope of this dissertation. We are interested in approaches based on energy minimization which provide an exciting way to solving imaging and vision problems and have been successfully applied to segmentation problems.

In general, the energy-based approaches can be classified into approaches based on a continuous or discrete setting. In the continuous setting, the methods are usually based on a variational approach. This includes active contours or snakes [83], geodesic active contour [27], geodesic active region [115], the Mumford-Shah model [106], region competition [155], active contour without edges [31], etc. In the discrete setting on the other hand the approaches are usually graph-based. This includes the well-known graph-cuts methods [17, 87].

The energy-based approaches formulate the segmentation problem into an optimization problem which can be expressed in several forms depending on the data and application. The idea of the approaches is that the segmentation process is performed by minimizing a given energy or cost functional with respect to the unknown of the problem whose optimal solution corresponds to the object of interest in the given image. The optimization problem can be formulated as

$$\arg\min_{u} E(u), \tag{1.1}$$

where $E(u)$ is an energy functional and u can be contours or an approximation of an image. This way of formulating the segmentation problems allows us to conveniently incorporate some prior knowledge, such as regularizers and shape priors, into the energy functionals in order to obtain satisfactory results.

Contour Representations

Contours and their evolutions can be represented using parametric or non-parametric models. In parametric representation models, the contours Γ are represented explicitly by using, for instance: a polygon [143] or a spline [13, 45], and driven by control points.

In non-parametric representation models, which are used in this dissertation, the contours are represented implicitly by higher dimensional functions. Although this way of representing the contours makes the evolution of the contours computationally more expensive, it has several advantages, such as: topological changes such as splitting and merging are handled automatically, discretization is done on a fixed grid, and the contour does not need to be reparameterized, avoiding numerical instabilities.

The level set methods, introduced by Dervieux and Thomasset [53] more than three decades ago and rediscovered a few years later by Osher and Sethian [111], have proven to be successful as numerical schemes for dealing with moving fronts and interfaces and have been widely used in the areas of image processing, computer vision and computer graphics. The idea is that a contour $\Gamma \subset \mathbb{R}^n$ is represented implicitly by the zero-level set of a higher dimensional Lipschitz function $\phi : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^n$ is an open region. The level set function ϕ is defined as follows:

$$\begin{cases} \phi(x) > 0 & \text{for } x \in \Omega_{int} \\ \phi(x) = 0 & \text{for } x \in \Gamma \\ \phi(x) < 0 & \text{for } x \in \Omega_{out}, \end{cases} \tag{1.2}$$

where Ω_{int} and Ω_{out} are, respectively, the region inside and outside Γ (see Figure 1.2). The geometric characteristics such as the unit normal and the mean curvature, and the motion of the contour can then be computed with this level set function. See [110] for more details.

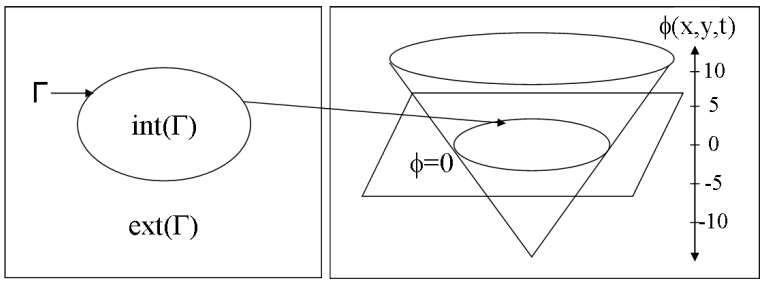


Figure 1.2: The curve Γ is represented by the zero-level set of the function ϕ.

Variational Approaches

Variational approaches are defined in a continuous setting and are mathematically well defined. There are many problems in computer vision and image analysis that can be formulated in a variational framework: Given an energy or cost functional which describes a specific task in computer vision, the energy is then minimized with respect to the unknowns of the problem to find an optimal solution. The variational models can be written as follows: For a given energy functional $E(u)$ with the unknown u

$$\inf\left\{E(u)=\int_{\Omega} f\big(x,u(x),\nabla u(x)\big)\,dx,\ u\in V\right\}, \tag{1.3}$$

the optimal solution u^*, defined in an appropriate space V, can be obtained by minimizing $E(u)$ such that

$$u^* = \arg\min_{u\in V} E(u). \tag{1.4}$$

The calculus of variations provides a framework for finding the minimum of a functional in which the fundamental theorem of this approach, the Euler-Lagrange equation, gives a necessary condition for u^* to be the minimum of $E(u)$ by computing the first variation if E is continuous and differentiable:

$$\frac{\partial E}{\partial u}=0. \tag{1.5}$$

The (local) minima can then be computed by using the gradient descent/ascent method until the steady state:

$$\frac{\partial u}{\partial t}=\pm\frac{\partial E}{\partial u}, \tag{1.6}$$

where t is an artificial time variable. The shape derivative tool, proposed by Delfour and Zolezio in [50], can also be used to find the Euler-Lagrange equation of the corresponding region-based energy functional, see [6] in which Aubert *et al.* described the use of calculus of variations and the shape derivative tool for variational image segmentation problems.

1.2.1 The Active Contour Models

Here we describe shortly some prominent active contour models (see [23] for more detail review of active contour models). The active contour models can be classified into edge-based and region-based models. The models originated by the snakes of Kass, Witkin and Terzepoulos [83] who proposed a spline-based parametric model that imposes the alignment of the

contour with the local minima defined by image gradients while maintaining its smoothness by minimizing an energy functional of the form

$$E_{Snake}(\Gamma) = \int_0^1 \left\{ \alpha|\Gamma'(p)|^2 + \beta|\Gamma''(p)|^2 - |\nabla I(\Gamma(p))| \right\} dp, \tag{1.7}$$

where α and β are some weight parameters, $\Gamma(p)$ denotes an explicit parametric contour and I is a given image. The first two terms are the regularization terms which impose the smoothness of the contour. The last term is the data term which attracts the contour toward the image gradients. This model is dependent on the chosen parametrization of the initial contour $\Gamma(p) = (x(p), y(p)) \in \Omega$, $p \in [0,1]$ and the contour Γ has to be initialized close to the object of interest.

Caselles *et al.* [27] proposed a geometrically intrinsic model of geodesic active contour (GAC). This model reformulates the segmentation as the problem of finding a contour of smallest length (geodesic) in a Riemannian space whose metric is induced by image gradients. The energy functional is given by

$$\begin{aligned} E_{GAC}(\Gamma) &= \int_0^1 g(|\nabla I(\Gamma(p))|)|\Gamma'(p)|\, dp \\ &= \int_0^{L(\Gamma)} g(|\nabla I(\Gamma(s))|)\, ds, \end{aligned} \tag{1.8}$$

where ds is the Euclidean length element, $L(\Gamma)$ is the Euclidean length of the contour Γ, defined by $L(\Gamma) = \int_0^1 |\Gamma'(p)|\, dp = \int_0^{L(\Gamma)} ds$ and g is an edge detector function. The energy of GAC E_{GAC}, which can be seen as a weighted length of the contour Γ, is independent on the parametrization of the contour. This implies that the model can be realized in the level set framework [111], as done in [27]. This is not the case for the Snakes model (1.7). GAC model was further improved by using zero-crossings of the image Laplacian edge detector which provides an optimal edge-integration [84].

The above models are examples of edge-based models. The edge-based models encounter problems if the objects or regions of interest have diffused edges, especially since the given image is usually presmoothed at a certain scale that can remove some image information at boundaries. This problem can be overcome by using region-based models that smooth the regions while preserving the boundaries. Most region-based segmentation models such as the works of Zhu and Yuille [155], Paragios and Deriche [115], Chan and Vese [31], Jehan-Besson *et al.* [77], etc., are influenced by the Mumford-Shah functional [105, 106]. The key assumption of the region-based models is that meaningful image regions are homogeneous in terms of

their intensities, color or texture.

The Mumford-Shah Functional

Mumford and Shah (MS) in [105] and [106] observed that a large class of images can be decomposed into homogeneous regions. They proposed to approximate a given image $I : \Omega \mapsto \mathbb{R}$ on a domain Ω with a piecewise smooth function u by minimizing the energy functional

$$E_{MS}(u,\Gamma) = \lambda \int_{\Omega} (I-u)^2\, dx + \nu \int_{\Omega\setminus\Gamma} |\nabla u|^2\, dx + |\Gamma|, \tag{1.9}$$

simultaneously with respect to u and the set of discontinuities Γ (the region boundaries). The first term is a data fidelity term which enforces the function u to be close to the given image, the second term is a regularizing term which smoothes the segmented image while preserving discontinuities of u across the boundaries defined by Γ and the last term is a regularizing term that penalizes the length of the region boundaries. λ and ν are weight parameters. This functional was also formulated in a discrete setting by Geman and Geman [69] using Markov Random Fields.

There have been several approaches proposed for solving the Mumford-Shah functional, for instance: A coarse to fine method proposed by Blake and Zisserman [14], Γ-convergence proposed by Ambrosio and Tortorelli [2], finite difference approximation by Gobbino [107], finite-element approximation by Bourdin and Chambolle [15], the level-set implementations by Chan and Vese [31], etc.

The Chan-Vese Model

Chan and Vese (CV) [31] proposed a two-phase level set-based active contour model to detect objects whose boundaries are not necessarily defined by image gradients. This is a particular case of the minimal partition problem of the Mumford and Shah functional [106]. The model approximates a given image $I : \Omega \mapsto \mathbb{R}$ by a piecewise constant instead of a piecewise smooth function where the active contour evolves by minimizing the difference between the intensity values of the two regions (regions inside and outside the contour) and mean intensity values inside and outside the contour, respectively. This can be done by minimizing the energy functional

$$E_{CV}(\Gamma,\mu_1,\mu_2) = \lambda \left\{ \int_{\text{int}(\Gamma)} (I-\mu_1)^2\, dx + \int_{\text{ext}(\Gamma)} (I-\mu_2)^2\, dx \right\} + |\Gamma|, \tag{1.10}$$

where μ_1, μ_2 are mean intensity values inside and outside the contour respectively, λ is a weight parameter, and $|\Gamma|$ is the length of the contour. The energy functional is then formulated using the level set method:

$$E_{CV}(\phi, \mu_1, \mu_2) = \lambda \left\{ \int_\Omega (I-\mu_1)^2 H(\phi)\, dx + \int_\Omega (I-\mu_2)^2 (1-H(\phi))\, dx \right\} + \int_\Omega \delta(\phi)|\nabla\phi|\, dx, \tag{1.11}$$

where $H(\phi)$ denotes the Heaviside function

$$H(t) = \begin{cases} 1, & t > 0, \\ 0, & t \leq 0, \end{cases} \tag{1.12}$$

and $\delta(\phi)$ denotes the delta function

$$\delta(t) = \begin{cases} +\infty, & t = 0, \\ 0, & \text{elsewhere}, \end{cases} \quad \text{s.t.} \int_{-\infty}^{\infty} \delta(t)\, dt = 1, \tag{1.13}$$

and the region mean intensity values can be computed in terms of the level set function by

$$\mu_1 = \frac{\int_\Omega I(x) H(\phi(x))\, dx}{\int_\Omega H(\phi(x))\, dx}, \tag{1.14}$$

$$\mu_2 = \frac{\int_\Omega I(x)(1-H(\phi(x)))\, dx}{\int_\Omega (1-H(\phi(x)))\, dx}. \tag{1.15}$$

Keeping μ_1, μ_2 fixed, this functional is minimized by gradient descent on ϕ

$$\frac{\partial \phi}{\partial t} = |\nabla\phi| \left[\operatorname{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right) - \lambda \left\{ (I-\mu_1)^2 - \lambda (I-\mu_2)^2 \right\} \right], \tag{1.16}$$
$$\phi(x,0) = \phi_0(x), \quad x \in \Omega,$$

where $t > 0$ is an artificial time variable. Figure 1.3 shows the segmentation of a synthetic image into two regions using the Chan-Vese model, where the initial contour (left) is evolved (middle) by solving (1.16) until the steady state to obtain the optimal solution (right).

The extension of the two-phase Chan-Vese model into a multi-phase one is straightforward. Vese and Chan (VC) [147] have proposed a multi-phase segmentation model using the Mumford-Shah model in a level set framework. In their work, in order to segment n regions, N level set functions $\{\phi_i\}_{i=1}^N$ are needed where $N = \log_2(n)$. For an illustration, assume that we would like to segment an image into four regions, see Figure 1.4. The energy functional (1.11)

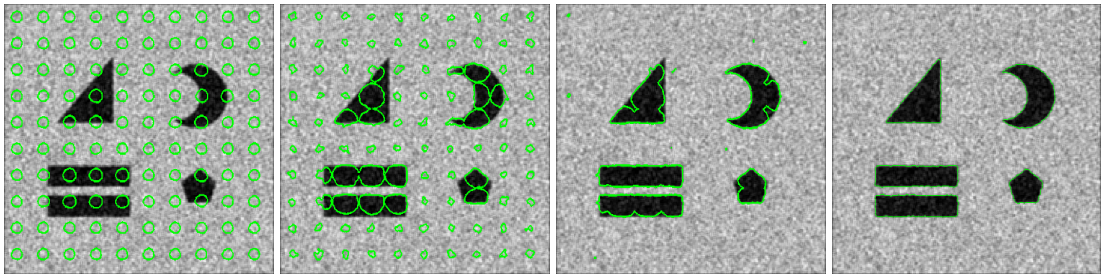


Figure 1.3: An example of contour evolution in image segmentation using the Chan-Vese model.

can then be formulated in the form:

$$\begin{aligned} E_{VC}(\Phi,\mu) =& \lambda \Bigg\{ \int_\Omega (I-\mu_{11})^2 H(\phi_1)H(\phi_2)\,dx + \int_\Omega (I-\mu_{10})^2 H(\phi_1)(1-H(\phi_2))\,dx \\ &+ \int_\Omega (I-\mu_{01})^2 (1-H(\phi_1))H(\phi_2)\,dx \\ &+ \int_\Omega (I-\mu_{00})^2 (1-H(\phi_1))(1-H(\phi_2))\,dx \Bigg\} \\ &+ \int_\Omega \delta(\phi_1)|\nabla\phi_1|\,dx + \int_\Omega \delta(\phi_2)|\nabla\phi_2|\,dx, \end{aligned} \tag{1.17}$$

where $\mu = (\mu_{11}, \mu_{10}, \mu_{01}, \mu_{00})$ is a constant vector of mean intensity values for each region, and $\Phi = (\phi_1, \phi_2)$ is a vector of two level set functions. See [147] for details.

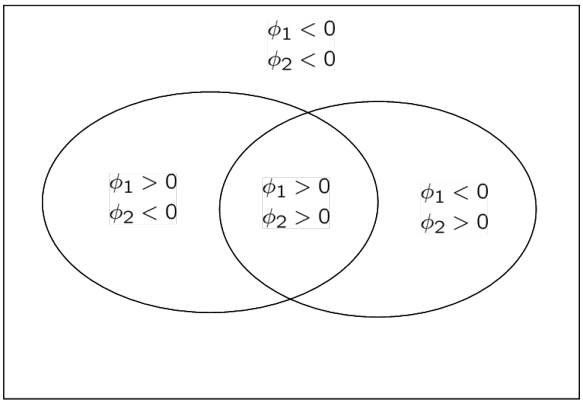


Figure 1.4: Two level set functions are needed to segment four regions using the multi-phase model of Vese and Chan [147].

1.2.2 Convex Formulation of the Active Contour Models

Most energy functionals of the active contour models are non-convex, which makes the initialization critical to obtain optimal solutions. The models are prone to get stuck in local minima since they are usually minimized by gradient descent methods. In practice, local minima can be avoided if the initialization is close enough to the desired solution.

In order to deal with the existence of local minima, Cohen and Kimmel in [36] proposed a minimal path approach to find a global minimum for the geodesic active contour model. Chan *et al.* [33] proposed another approach. The idea is to convexify the energy functional of the piecewise constant Mumford-Shah functional [106] of Chan and Vese [31] by extending the Chan-Vese functional (1.10) and its minimization in such a way that the result can be transformed into a global minimizer of the original problem by simple thresholding. This method is based on observations of Strang [141] on maximal flows through a continuous domain in which the coarea formula of Fleming and Rishel [60] is the primary tool.

The idea is the following: Let $I : \Omega \mapsto \mathbb{R}$ be an input image where $\Omega \subset \mathbb{R}^n$ is the image domain and consider the following energy functional

$$E(\Sigma, \mu_1, \mu_2) = \mathrm{Per}(\Sigma; \Omega) + \lambda \left\{ \int_\Sigma (I(x) - \mu_1)^2 \, dx + \int_{\Omega \setminus \Sigma} (I(x) - \mu_2)^2 \, dx \right\}, \tag{1.18}$$

where λ is a weight parameter, $\Sigma \subset \Omega$ denotes a subset of Ω, and $\mathrm{Per}(\Sigma; \Omega)$ denotes the perimeter of Σ. The segmentation is carried out by solving the following optimization problem

$$\min_{\mu_1, \mu_2 \in \mathbb{R}; \Sigma \subset \Omega} E(\Sigma, \mu_1, \mu_2), \tag{1.19}$$

which looks for the best approximation to the image $I(x)$ in the L^2 sense among all functions that take only two values μ_1, μ_2. The problem of (1.18) is solved by two alternating steps where in the first step μ_1, μ_2 are computed for fixed Σ using the usual formula

$$\mu_1 = \frac{1}{|\Sigma|} \int_\Sigma I(x) \, dx \quad \text{and} \quad \mu_2 = \frac{1}{|\Omega \setminus \Sigma|} \int_{\Omega \setminus \Sigma} I(x) \, dx, \tag{1.20}$$

and the second step updates the Σ for fixed μ_1, μ_2.

As pointed out by Chan *et al.* [33], if $\mu_1, \mu_2 \in \{0, 1\}$ and $I(x)$ is taken to be the characteristic function $\mathbf{1}_\Omega(x)$, then the minimization problem of 1.19 reduces to the image denoising model of Rudin, Osher and Fatemi [132] for binary functions.

In the level set framework of Chan and Vese [31], the boundary of Σ, i.e. $\partial\Sigma$, is represented by the zero-level set of function ϕ. The energy functional (1.18) can then be written in the level set framework as

$$E_{CV}(\phi,\mu_1,\mu_2) = \int_\Omega |\nabla H_\varepsilon(\phi)|\,dx + \lambda \int_\Omega H_\varepsilon(\phi)(I(x)-\mu_1)^2 + (1-H_\varepsilon(\phi))(I(x)-\mu_2)^2\,dx, \tag{1.21}$$

where $\lambda \in \mathbb{R}$ and the function H_ε is a regularized Heaviside function. The associated Euler-Lagrange equation is given by

$$\frac{\partial\phi}{\partial t} = |\nabla\phi|\left[\operatorname{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right) - \lambda\left\{(I(x)-\mu_1)^2 - (I(x)-\mu_2)^2\right\}\right]. \tag{1.22}$$

The equation (1.22) has the same steady state solutions as:

$$\frac{\partial\phi}{\partial t} = \operatorname{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right) - \lambda\left\{(I(x)-\mu_1)^2 - (I(x)-\mu_2)^2\right\}, \tag{1.23}$$

which is the Euler-Lagrange equation of the following energy functional:

$$\int_\Omega |\nabla\phi(x)|\,dx + \lambda \int_\Omega \left\{(I(x)-\mu_1)^2 - (I(x)-\mu_2)^2\right\}\phi(x)\,dx. \tag{1.24}$$

The energy functional (1.24) does not have a minimizer in general since it is linear in ϕ. By carrying out the gradient descent of (1.23), the level set function ϕ approaches $\pm\infty$ depending on the sign of ϕ due to the nonuniqueness of representation with level sets. To get rid of this, the function ϕ is restricted such that $0 \le \phi(x) \le 1$ for all $x \in \Omega$.

As proposed and proven by Chan *et al.* [33], by introducing an auxiliary variable u, and for fixed $\mu_1,\mu_2 \in \mathbb{R}$, a global minimizer for the energy functional (1.18) can be found by solving the following convex minimization problem

$$\min_{0\le u\le 1} \int_\Omega |\nabla u(x)|\,dx + \lambda \int_\Omega \left\{(I(x)-\mu_1)^2 - (I(x)-\mu_2)^2\right\}u(x)\,dx, \tag{1.25}$$

by setting $\Sigma = \{x : u(x) > \gamma\}$ for almost any choice of $\gamma \in [0,1]$.

Another version of (1.25) was proposed by Mory and Ardon [104], referred to as the *fuzzy region competition* for convex two-phase segmentation by solving

$$\min_{0\le u\le 1}\left\{\int_\Omega |\nabla u(x)|\,dx + \lambda \int_\Omega \left\{u(x)(I(x)-\mu_1)^2 + (1-u(x))(I(x)-\mu_2)^2\right\}dx\right\}, \tag{1.26}$$

where

$$\mu_1 = \frac{\int_\Omega u(x)I(x)\,dx}{\int_\Omega u(x)\,dx} \quad \text{and} \quad \mu_2 = \frac{\int_\Omega (1-u(x))I(x)\,dx}{\int_\Omega (1-u(x))\,dx}. \tag{1.27}$$

See [33] and [104] for details. By having a convex functional, a global solution can be obtained, even by using local optimization methods such as gradient descent. Figure 1.5 shows the comparison of the segmentation of a horse using the same value of λ between the original Chan-Vese model (1.11) and convex formulation of the Chan-Vese model (1.26). In Chapter 4 we will extend the convex formulation of the Chan-Vese model (1.25) by adding an additional shape prior term in order to segment an object of interest in an image.

Figure 1.5: Segmentation of a horse with the Chan-Vese model (left) and convex formulation of the Chan-Vese model (right).

In the past three years, since the convex formulation of the Mumford-Shah model was introduced by Chan *et al.* [33], some work has been done to extend the Mumford-Shah segmentation models from two-region into multi-region problems such as in Pock *et al.* [121], Zach *et al.* [152], Lellmann *et al.* [92], Chambolle *et al.* [30], Pock *et al.* [120], Berkels [11], Olsson *et al.* [108], Brown *et al.* [24], etc. In the simplest form, the piecewise constant Mumford-Shah model of Chan and Vese, can be regarded as a continuous version of the Potts model [125]. This model tries to partition the continuous domain Ω into n subdomains $\{\Omega_i\}_{i=1}^n$ by minimizing the energy functional

$$E(\{\Omega_i\}_{i=1}^n) = \lambda \sum_{i=1}^{n} \int_{\Omega_i} (I(x)-\mu_i)^2\,dx + \sum_{i=1}^{n} |\partial\Omega_i|, \tag{1.28}$$

where $I(x)$ is an image, $|\partial\Omega_i|$ measures the lengths of the boundaries of the disjoint subdomains $\Omega_i, i = 1,...,n$ and $\mu_i, i = 1,...,n$ are constants. In order to obtain the global minimum of the

energy functional (1.28), Zach *et al.* in [152] proposes the following energy functional

$$E(u) = \int_{\Omega} \left(\lambda \sum_{i=1}^{n} (I(x) - \mu_i)^2 u_i + \sum_{i=1}^{n} |\nabla u_i| \right) dx, \tag{1.29}$$

with simplex constraints $u_i \geq 0$ and $\sum_{i=1}^{n} u_i = 1, i = 1, ..., n$. A similar formulation with slightly different regularization, but with the same simplex constraints, is also proposed by Lellmann *et al.* [92]:

$$E(u) = \int_{\Omega} \left(\lambda \sum_{i=1}^{n} (I(x) - \mu_i)^2 u_i + \sqrt{\sum_{i=1}^{n} |\nabla u_i|^2} \right) dx. \tag{1.30}$$

We will use the energy functional of (1.30) and adapt it to the multi-region segmentation problems on the manifold in Chapter 7.

1.3 Prior Knowledge in Segmentation Problems

Since segmentation problems are generally ill-posed, unconstrained segmentation models give unsatisfactory results, particularly in the presence of noise, clutter and occlusions. Therefore, prior knowledge is needed to constrain the segmentation process in order to extract meaningful regions or objects in an image.

In general, we would like to segment a given image I by finding a contour Γ that separates an object and its background by minimizing an energy functional of the form:

$$E(\Gamma) = E_{Data}(I, \Gamma) + \lambda E_{Prior}(\Gamma). \tag{1.31}$$

The first term on the right hand side is the data fidelity term and the second term is the prior knowledge about the objects or regions of interest, and λ is a weight parameter which controls the influence of the prior knowledge in the segmentation process.

This energy functional can also be modeled in a Bayesian framework as in [155]. In order to find a contour Γ, one maximizes the posterior probability of the form:

$$P(\Gamma|I) \propto P(I|\Gamma) P(\Gamma), \tag{1.32}$$

which is equivalent to minimizing its negative log-likelihood

$$-\log P(\Gamma|I) \propto -\log P(I|\Gamma) - \log P(\Gamma), \tag{1.33}$$

such that $-\log P(I|\Gamma) = E_{Data}(I,\Gamma)$ and $-\log P(\Gamma) = E_{Prior}(\Gamma)$. This shows that the data fidelity term is related to the probability of the intensity distribution I given a contour Γ and the prior knowledge term is related to the a priori probability for a given contour Γ.

The characteristic of prior knowledge needed to segment an object of interest in a given image varies depending on the features that distinguish the object from its background. Local priors such as smoothness constraints and global priors such as shape constraints can be used in the segmentation process. While local models using intensity and curvature models such as in [98] have been successfully applied to segmentation problems, global models are more robust to severe occlusion which is one of the most difficult problems to deal with.

1.3.1 Total Variation

In most active contour models in Section 1.2, the prior term is a local smoothness term $E_{Prior}(\Gamma) = |\Gamma|$ which constrains the contour Γ to be as smooth as possible. A particularly interesting prior is the smoothness term of the Total Variation (TV) functional of a function u:

$$TV(u) = \int_{\Omega} |\nabla u| \, dx, \tag{1.34}$$

where

$$|\nabla u| = \sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2}. \tag{1.35}$$

The TV functional has been used in many applications related to image processing. It was introduced as a PDE-based model for edge-preserving noise removal by Rudin, Osher and Fatemi (ROF) in [132], where it removes small oscillations such as noise but preserves sharp discontinuities such as edges. The idea of this model is to denoise a noisy image I by using the TV functional (1.34) as a regularization term and with an L_2 fidelity term in the ROF functional

$$E_{ROF}(u) = \int_{\Omega} |\nabla u| \, dx + \lambda \int_{\Omega} (u-I)^2 \, dx. \tag{1.36}$$

$TV(u)$ is finite for any bounded functions, even for discontinuous functions as long as u does not have infinitely large jumps. $TV(u)$ can be expressed in terms of the integral the level sets of u by using the coarea formula of Fleming and Risher [60]

$$\int_{\Omega} |\nabla u| \, dx = \int_{-\infty}^{\infty} Per(\{x : u(x) > \gamma\}) \, d\gamma, \tag{1.37}$$

where one integrates the perimeter of the set $\{x : u(x) > \gamma\}$ for all values of γ. This formula is used to express energy functionals and the constraints in terms of super level sets in Subsection 1.2.2. See [33, 122] for more details and interesting applications of the TV functional in Image Analysis and Computer Vision.

1.3.2 Shape Priors

Enforcing a prior knowledge about the shape of objects is a natural way to help the segmentation process to segment meaningful objects from images, especially in the presence of noise which results in spurious boundaries, and of occlusions that cause overlaps of object boundaries. In fact, these situations always happen in real applications. However, the integration of global shape constraints into the segmentation process is non-trivial. Consequently, when dealing with shape priors for segmentation problems, we need to start defining the shape, the set of admissible shapes and its properties, and a way to compare shapes, that is a distance measure between two shapes. There are several ways to enforce shape constraints in a segmentation process using active contour models. Many approaches incorporate an additional shape prior term into the segmentation energy functional

$$E(\Gamma) = E_{Data}(I,\Gamma) + \lambda_1 E_{Smoothness}(\Gamma) + \lambda_2 E_{Shape}(\Gamma,\Gamma_0), \tag{1.38}$$

where Γ_0 is a template of the shape of interest and $E_{Shape}(\Gamma,\Gamma_0)$ measures the distance between Γ and Γ_0. This way of integrating the shape knowledge has been used in some work, such as Cremers *et al.* [39, 45], Chen *et al.* [35], Rousson and Paragios [128, 129], Zang and Freedman [154], Cremers and Soatto [42], Rousson and Cremers [130], Chan and Zhu [32], Bresson *et al.* [22], Foulonneau *et al.* [61], and references therein. The cumbersome part of this approach is having to adjust the parameters λ_1, λ_2 to balance the fidelity term E_{Data}, the smoothness term $E_{Smoothness}$ and the shape prior term E_{Shape} according to the level of noise, clutter and occlusions, and to the confidence in the model.

The shape prior term E_{Shape}, which measures the distance between the evolving curve Γ and a reference shape Γ_0 is naturally defined by

$$E_{Shape}(\Gamma,\Gamma_0) = \int_0^1 \Big(\Gamma(s) - \Gamma_0(s)\Big)^2 ds. \tag{1.39}$$

Since we are interested in using implicit representations of the contours, the idea of (1.39) can be used in terms of level set functions where the contours Γ and Γ_0 are the zero-level set of ϕ

and ϕ_0 respectively, such that we get a new energy functional

$$E_{Shape}(\phi, \phi_0) = \int_\Omega \Big(\phi(x) - \phi_0(x)\Big)^2 dx. \tag{1.40}$$

The reference shape ϕ_0 can be a template of a silhouette such as in [126] or a shape obtained from a shape model which is built using some learning procedure [22, 35, 40, 41, 59, 93, 128, 146] or a result from previous a segmentation process especially in tracking applications such as in [103, 154]. The three possibilities are used in the work of this dissertation. Figure 1.6 is an example of the segmentation of a horse using the Chan-Vese model with an additional shape prior term.

It is important that the variation of reference shape is not far away from the shape being segmented in order to obtain optimal segmentation results. This makes the shape variability due to the non-rigidity of the object, noise, occlusions, the changes in the object pose, illumination, camera's view point, and etc., one of the most important issues to deal with. This issue is mainly handled by using statistical models based on Principal Component Analysis (PCA) that can model global variations samples within the training set, such as in [21, 37, 93, 146] or by using more complex models such as in [40, 41, 46, 58].

Pose Invariance

Objects in the image can have different scales, orientation, etc compared to the shape of reference. This causes one of the main difficulties in dealing with shape prior based models, that is, the need to align the shape of reference to the shape being segmented such that the shape of reference is invariant with respect to a group of transformations. Pose parameters such as scaling, rotation and translation can be taken into account either in explicit or implicit setting. In an explicit setting such as in [35, 93, 126, 128, 146], the pose parameters are optimized during the segmentation process, which makes the problem more complex. The shape prior term (1.40) can then be written as

$$E_{Shape}(s, R, \mathbf{a}, \phi, \phi_0) = \int_\Omega \Big(\phi(x) - \phi_0(sRx + \mathbf{a})\Big)^2 dx, \tag{1.41}$$

where s is a scaling parameter, R is a rotation matrix, and $\mathbf{a}$ is a translation vector.

As noted by Cremers *et al.* [41], optimization to obtain optimal transformation parameters explicitly has several drawbacks. When the optimization is done by using gradient descent

method, one needs a careful tuning process, in numerical experiments, to determine an appropriate time step for each parameter to guarantee the stability. It is also unclear how to alternate between the updates of the respective parameters and how often one should iterate each gradient descent equation. To overcome those problems, an implicit setting can be used, such as using intrinsic alignment proposed by Cremers *et al.* [41], moment-based shape priors of Foulonneau *et al.* [61], etc. In this dissertation we use the explicit setting due to the simplicity in deriving the gradient descent equations and in their implementations and propose a way to solve the time step issue in Chapter 6.

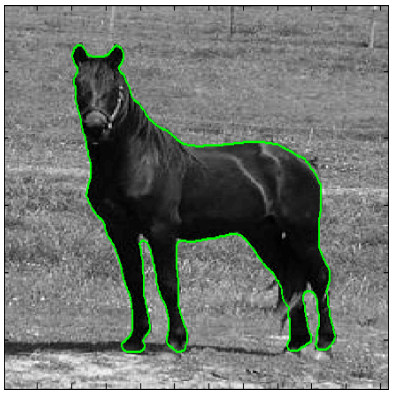


Figure 1.6: Segmentation of a horse in Figure 1.5 is improved by adding a shape prior term.

1.4 Contributions/Summaries of the Papers

Motivated by the strong mathematical foundation, being independent of grid-bias and having a convenient way to integrate prior knowledge, especially the shape priors, we use variational segmentation models, instead of graph-based models, for solving segmentation problems. Other motivations of this choice that are the recent progresses in the convex formulation of the Mumford-Shah models and the fast numerical schemes for solving TV-based models such as in [20, 72] and the references therein. In this dissertation, we propose region-based active contour models of Chan and Vese with additional prior knowledge, such as occlusion information, shape knowledge and the total variation as the smoothness term, in order to improve the robustness of the models in the presence of noise, clutter and occlusions.

The content of this dissertation is based on the result of joint work with Niels Chr. Overgaard, David Gustavssson, Christian Gosch, Amael Delaunoy, Emmanuel Prados, Christoph

Schnoerr, Mads Nielsen and Anders Heyden, which has been presented in [48, 64, 65, 66, 67, 68, 73, 74, 112]. We summarize the main contributions of the dissertation as follows:

Variational Segmentation of Image Sequences using Region-Based Active Contours and Deformable Shape Priors

Chapter 2 is a reformatted and slightly modified version of paper [65] without affecting the content or the results. Here we address the problem of segmentation of nonrigid, moving objects in image sequences in the presence of noise, clutter and occlusions by using region-based active contours of Chan and Vese [31] augmented with a frame-to-frame interaction term as a shape prior. In order to take advantage of the prior knowledge obtained from the segmentation of previous frames, we propose to use the interaction term $E_I = E_I(\Gamma, \Gamma_0)$ which penalizes deviations of the current active contour Γ from the previous one, Γ_0. Instead of using the shape prior energy term as in (1.41), we formulate our shape energy term as

$$E_I(\Gamma, \Gamma_0) = \min_T \int_{\mathrm{int}(\Gamma)} \phi_0(T^{-1}x)\, dx, \tag{1.42}$$

where $\phi_0 : \Omega \to \mathbb{R}$ denotes the level set function of the contour Γ_0, and the minimum is taken over the *group of Euclidean transformations* $T : \mathbb{R}^2 \to \mathbb{R}^2$ which preserves the orientation of the plane. The benefit of our interaction term E_I is that its L^2-gradient can be computed easily by

$$\nabla_\Gamma E_I(\Gamma, \Gamma_0) = \phi_0(x) = \phi(\Gamma_0; x) \quad (x \in \Gamma),$$

and that this gradient is small if Γ is close to the shape prior Γ_0, and large if the active contour is far from the shape prior. The prior-segmentation model can then be formulated as the following energy functional

$$E(\Gamma) = E_{CV}(\Gamma) + \lambda E_I(\Gamma, \Gamma_0), \tag{1.43}$$

where E_{CV} is a reduced Chan-Vese functional (1.10), and $\lambda > 0$ is a coupling constant which determines the strength of the interaction. The energy (1.43) is then minimized using the gradient descent method in a level set framework for each frame in the image sequence.

The performance of the model is illustrated with experiments on synthetic and real image sequences.

Nonrigid Variational Object Segmentation with Occlusion Detection in Image Sequences

Chapter 3 is the result of merging papers [68, 74] with minor modifications without affecting the content or the results. We deal with the problem of segmentation of nonrigid objects that are partially occluded in image sequences using the segmentation method in Chapter 2 coupled with a variational contour matching formulation between two consecutive contours as a registration using the geometry-constrained diffusion equation [3]. The presence of occlusions, such as overlapping objects or the superposition of an object to another object, is one of the major problems in the segmentation of image sequences. While minor occlusions can usually be handled by using shape priors, severe occlusion is still a big problem. As formulated in (1.43) we need to tune the coupling constant λ to define the strength of the shape prior influence in the segmentation, which depends on the data and applications. In the presence of severe occlusions, we need to set λ high enough to be able to segment the object of interest, but this can cause the segmentation of the current frame to be very close to the previous one, which limits the capability of the model to capture object deformations. To overcome this problem, it is necessary to detect and locate the occlusions. If we have information about the occlusions, we can either use a spatially and temporally adaptive λ, or reconstruct the occluded region to improve the robustness of the segmentation models.

To detect and locate the occlusions, suppose we have two closed contours Γ_1 and Γ_2 in the image domain Ω as the results of segmentation using (1.43). Let $\Phi = \Phi(x) : \Omega \to \mathbb{R}^2$ such that $\Phi(\Gamma_1) = \Gamma_2$ is a mapping that can be expressed in the form $\Phi(x) = x + U(x)$, where the vector valued function $U = U(x) = (u_1(x), u_2(x))^T : \Omega \to \mathbb{R}^2$ is called the displacement field associated with Φ. We can then find the optimal mapping by minimizing the energy functional

$$E[\Phi] = \frac{1}{2}\int_{\Omega} |\nabla u_1(x)|^2 + |\nabla u_2(x)|^2 \, dx. \tag{1.44}$$

The mapping $\Phi(x)$ is an estimation of the displacement of the boundary of an object between two frames. By finding the displacement of the contour, a consistent displacement of the intensities inside the closed curve Γ_1 can also be found. Φ maps Γ_1 onto Γ_2 and pixels inside Γ_1 are mapped inside Γ_2. This displacement field which only depends on displacement of the contour, can then be used to map the intensities inside Γ_1 onto Γ_2. After the mapping, the intensities inside Γ_1 and Γ_2 can be compared and classified as the same or different value. After

the occlusions are detected and located using deviations from predicted intensities, the missing intensities in the occluded regions can be reconstructed. After reconstructing the occluded regions in the novel image, the segmentation can then be improved. Experimental results on synthetic and real image sequences are shown.

Continuous Cuts for Prior-Based Image Segmentation

Chapter 4 is a reformatted and slightly modified version of paper [64] without affecting the content or the results. We propose a novel prior-based variational object segmentation method in a global minimization framework of [33] which unifies image segmentation and image denoising. The idea of the proposed method is to extend the convex minimization problem (1.25) by adding an additional shape constraint in order to segment an object of interest in a given image. Inspired by the shape prior energy term of (1.41) and the fact that $\nabla_\phi E_{Shape}(s,R,\mathbf{a},\phi) = 2(\phi(x) - \phi_0(sRx+\mathbf{a}))$, the energy functional in the convex problem of (1.25) can then be reformulated by adding a shape prior term E_{Shape} such that

$$\min_{0\leq u\leq 1} \int_\Omega |\nabla u|\, dx + \lambda \int_\Omega \Big\{ (I(x)-\mu_1)^2 - (I(x)-\mu_2)^2 + \big(\hat{u} - \mathbf{1}_{\Omega_p}(sRx+\mathbf{a})\big) \Big\} u(x)\, dx, \quad (1.45)$$

where $\mathbf{1}_{\Omega_p}$ is the characteristic function of a shape prior template, $\hat{u}$ is a 'frozen' u which is updated after finding a solution to (1.45), and $s, R, \mathbf{a}$ are a scaling parameter, a rotation matrix and a translation vector. Experimental results demonstrate the performance and robustness of the method to segment objects of interest in real images.

View Point Tracking of Rigid Objects using Region-Based Segmentation Model and Shape Priors from Shape Submanifolds

Chapter 5 is a reformatted and slightly modified version of paper [73] without affecting the content or the results. We propose to use the region-based image segmentation models of [31] and [33] with shape priors from the shape submanifold of [101] to segment an object of interest in a sequence of images and use the object contours to infer and to track the viewpoint onto a 3D rigid object.

Assume we have a collection of silhouettes of a known object that corresponds to the object from different view point. These shapes can be regarded as samples of an object-specific submanifold of the manifold of all planar shapes that is parameterized by the view sphere. Taking into account the geometry of this submanifold and interpolating the shape samples

accordingly, the viewpoint of a moving camera, or object pose relative to the observer can be tracked by segmenting a given sequence $\{I_1,\dots,I_n\}$ of n images depicting a moving object by using the Chan-Vese model in (1.11) coupled with a shape prior energy based on [35] and [127] in the level set framework, but allowing only for Euclidean transformations:

$$E_{Shape}(s,R,\mathbf{a},\phi,\phi_0)=\int_\Omega \Big(H(\phi(x))-H(\phi_0(sRx+\mathbf{a}))\Big)^2\,dx, \tag{1.46}$$

where ϕ_0 is the shape prior template. As an alternative, we also use the convex segmentation model (1.45). The contour Γ_i, represented as the zero-level set of function ϕ_0 from the previous segmentation, is used for initialization and as a weak prior for the segmentation of image I_{i+1}. This approach replaces explicit 3D object models by the corresponding invariant shape submanifolds that are learnt from a sufficiently large number of image contours, and is applicable to arbitrary objects.

Pose Invariant Shape Prior Segmentation using Continuous Cuts and Gradient Descent on Lie Groups

Chapter 6 is a reformatted and slightly modified version of paper [112] without affecting the content or the results. We consider segmentation with pose invariant shape priors using continuous cuts as an extension of Chapter 4. The shape prior energy term is based on the L^2 shape dissimilarity measure and with pose invariance under the full (Lie-) group of similarity transforms in the plane. We consider shape priors f which are generated from a single shape template f_0 by similarity transforms $T=T(\theta,\sigma,\mathbf{a})$:

$$f(x)=f_0(T^{-1}(x)); \qquad T(x)=\mathbf{a}+e^{\sigma}\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}x. \tag{1.47}$$

The prior energy is then defined by

$$E_{Prior}(u)=\min_T \int_\Omega (u(x)-f_0\circ T^{-1}(x))^2\,dx. \tag{1.48}$$

The minimum of (1.48) is found by gradient descent in the (Lie-) group of similarity transformations. This requires the derivatives:

$$\begin{aligned} \frac{\partial}{\partial \mathbf{a}} E(\theta,\sigma,\mathbf{a}) = -\langle f-u, \nabla_x f\rangle, \quad \frac{\partial}{\partial \theta} E(\theta,\sigma,\mathbf{a}) = -\langle f-u, \nabla_x f^T J(\cdot-\mathbf{a})\rangle, \\ \text{and} \quad \frac{\partial}{\partial \sigma} E(\theta,\sigma,\mathbf{a}) = -\langle f-u, \nabla_x f^T (\cdot-\mathbf{a})\rangle, \end{aligned} \tag{1.49}$$

where $J = R(-\theta)^T R'(-\theta) = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$, to be computed. Notice that the x-derivatives are computed on the transformed prior $f = f_0 \circ T^{-1}$ - not on the template f_0.

To overcome the common numerical problems associated with step size control for translation, rotation and scaling in the discretization of the pose model, a new gradient descent procedure for the pose estimation is proposed based on the construction of a Riemannian structure on the group of transformations and a derivation of the corresponding pose energy gradient. The gradient descent differential equation for the pose transform $T = T(\theta,\sigma,\mathbf{a})$ is

$$\frac{d}{dt} T(t) = -\nabla E(T(t)), \qquad T(0) = T_0, \tag{1.50}$$

where the corresponding gradient of E_{Prior}, $\nabla E = (\nabla_\theta E, \nabla_\sigma E, \nabla_a E)$, has the components:

$$\begin{aligned} \nabla_{\mathbf{a}} E = \frac{\langle f-u, -\nabla_x f\rangle}{\|\nabla_x f\|^2}, \quad \nabla_\theta E = \frac{\langle f-u, -\nabla_x f^T J(\cdot-\mathbf{a})\rangle}{\| \, |x-\mathbf{a}| \nabla_x f \, \|^2}, \\ \text{and} \quad \nabla_\sigma E = \frac{\langle f-u, -\nabla_x f^T (\cdot-\mathbf{a})\rangle}{\| \, |x-\mathbf{a}| \nabla_x f \, \|^2}, \end{aligned} \tag{1.51}$$

which is used in the implementation of gradient descent search for the optimal pose parameters. The use of this gradient in the descent gives an automatic step-size control. Also, it allows us to use a single step-size which will work simultaneously for all four variables used to parameterize the group. Together with efficient numerics for TV-minimization such as in [28], we get a fast and reliable implementation of the model.

Convex Multi-Region Segmentation on Manifolds

Chapter 7 is a reformatted and slightly modified version of paper [48] without affecting the content or the results. Here we propose to adapt the convex segmentation model of (1.30) in order to segment data defined on a manifold into a set of regions with uniform properties and show how to optimize the energy for a manifold represented by triangular meshes.

First we rewrite the convex multi-region model in (1.30) using a general data term $\mathbf{s}(\mathbf{x})$:

$$\min_{\mathbf{u}\in K}\left\{\int_\Omega <\mathbf{u},\mathbf{s}> dx + \lambda \int_\Omega |\nabla \mathbf{u}|\, dx\right\}, \tag{1.52}$$

where K is a simplex constraint, i.e. the set of functions $\mathbf{u}:\Omega\rightarrow\mathbb{R}^m$ such that for all $\mathbf{x}\in\Omega$ and $i\in[1..m]$, $\mathbf{u}_i(\mathbf{x})\geq 0$ and $\sum_{i=1}^m \mathbf{u}_i(\mathbf{x})=1$; $|\nabla\mathbf{u}(\mathbf{x})|$ is the total variation of u that corresponds to $\sqrt{\sum_i |\nabla\mathbf{u}_i(\mathbf{x})|^2}$, where $|.|$ denotes the L^2 norm; m denotes the number of regions and $\mathbf{s}(\mathbf{x})$ is an m-dimensional vector, while $s_i(\mathbf{x})$ indicates the affinity of the data at point $\mathbf{x}$ with region i. The energy functional of (1.52) can then be formulated as follows:

$$\min_{\mathbf{u}\in K}\left\{\int_{\mathcal{S}} <\mathbf{u},\mathbf{s}> d\sigma + \int_{\mathcal{S}} \lambda |\nabla_{\mathcal{S}}\mathbf{u}|\, d\sigma\right\}, \tag{1.53}$$

where now the vector-valued function $\mathbf{u}$ is defined on a Riemannian manifold $\mathcal{S}$ instead of Ω, $|.|$ is the Riemannian norm, $\nabla_{\mathcal{S}}$ is the intrinsic gradient on $\mathcal{S}$ and $d\sigma$ is the manifold's element measure. Consider a manifold represented by triangular meshes, with $\mathcal{S}_j$ as the j^{th} triangle of the mesh and $\mathbf{U}=\{\mathbf{u}_k\}$ a discrete relaxation variable on the mesh, then the energy (1.53) can be rewritten as

$$\sum_j\sum_k \left\langle \mathbf{u}_k, \int_{\mathcal{S}_j}\phi_k \mathbf{s}\, d\sigma\right\rangle + \lambda\int_{\mathcal{S}_j}|\nabla_{\mathcal{S}}\mathbf{u}|d\sigma, \tag{1.54}$$

where $\phi_k:\mathcal{S}\rightarrow\mathbb{R}$ is the piecewise affine, interpolating basis function such that $\phi_k(\mathbf{x}_k)=1$ and $\phi_k(\mathbf{x}_i)=0$ if $i\neq k$. The total variation term is formulated explicitly using fundamental forms. The choice of the data term $\mathbf{s}$ depends on the application. For segmentation of a given image I using the piecewise constant Mumford-Shah model of Chan and Vese [31], we define $\mathbf{s}(\mathbf{x})=(I(\mathbf{x})-\mu)^2$ where μ are the mean values of the regions. In another example, such as the radiance segmentation from multiple views in 3D reconstruction problems, we can define $\mathbf{s}(\mathbf{x})=(I\circ\pi(\mathbf{x})-\mu)^2$ where π is the projection of $\mathbf{x}$ on image I.

Chapter 2

Variational Segmentation of Image Sequences using Region-Based Active Contours and Deformable Shape Priors

This chapter is a reformatted and slightly modified version of paper [65] without affecting the content or the results. In this chapter we address the problem of segmentation in image sequences using region-based active contours and level set methods. We propose a novel method for variational segmentation of image sequences containing nonrigid, moving objects. The method is based on the classical Chan-Vese model augmented with a novel frame-to-frame interaction term, which allow us to update the segmentation result from one image frame to the next using the previous segmentation result as a shape prior. The interaction term is constructed to be pose-invariant and to allow moderate deformations in shape. It is expected to handle the appearance of occlusions which otherwise can make segmentation fail. The performance of the model is illustrated with experiments on synthetic and real image sequences.

2.1 Introduction

The segmentation of objects in image sequences is an important and difficult problem in computer vision with applications to e.g. video surveillance. This is a difficult process in computer

vision, with the purpose of dividing a given image into one or several meaningful regions or objects. This process is more difficult when the objects to be segmented are moving and nonrigid. The shape of nonrigid, moving objects may vary a lot along image sequences due to, for instance, deformations or occlusions, which puts additional constraints on the segmentation process.

There have been a number of methods proposed and applied to the segmentation problems. Active contours are powerful methods for image segmentation; either boundary-based such as snakes [83] and geodesic active contours [27], or region-based such as geodesic active region [116] and Chan-Vese models [31], which are formulated as variational problems. Those variational formulations perform quite well and have often been applied using level sets. Active contour based segmentation methods often fail due to noise, clutter and occlusion. In order to make the segmentation process robust against these effects, shape priors have been proposed to be incorporated into the segmentation process. In recent years, many researchers have successfully introduced shape priors into segmentation methods such as in [22, 32, 35, 38, 41, 42, 44, 93, 128, 129, 130, 144, 146] and references therein.

We are interested in segmenting nonrigid moving objects in image sequences. When the objects are nonrigid, an appropriate segmentation method that can deal with shape deformations should be used. The application of active contour methods for segmentation in image sequences gives promising results. In [70], a snake-based segmentation is proposed, specifically designed to improve robustness against occlusion in the context of tracking. In [103, 114, 116], variants of the classical Chan-Vese model are used as the basis for segmentation. In [103], for instance, it is proposed to simply use the result from one image as an initializer in the segmentation of the next.

The main purpose of this chapter is to propose and analyze a novel variational segmentation method for image sequences, that can both deal with shape deformations and at the same time is robust to noise, clutter and occlusions. The proposed method is based on minimizing an energy functional containing the classical Chan-Vese functional as one part and a term that penalizes the deviation of the shape being segmented from the previous shape as a second part. The second part of the functional is based on a transformed distance map to the previous contour, where different transformation groups, such as Euclidean, similarity or affine, can be used depending on the particular application.

2.2 Theoretical Background

2.2.1 Region-Based Segmentation

We begin with a brief review of the classical Chan-Vese segmentation model [31]. In this model a gray scale image is considered to be a real valued function $I : \Omega \to \mathbb{R}$ defined on the *image domain* $\Omega \subset \mathbb{R}^2$, usually a rectangle. A point $x \in \Omega$ is often referred to as a pixel, and the function value $I = I(x)$ as the *pixel value*, or the *gray scale value*. The Chan-Vese model is an active contour model. The idea is to find a contour Γ, by which we mean a finite union of disjoint, simple, closed curves in Ω, such that the image I is optimally approximated by a single gray scale value μ_{int} on $\text{int}(\Gamma)$, the *inside* of Γ, and by another gray scale value μ_{ext} on $\text{ext}(\Gamma)$, the *outside* of Γ. The optimal contour Γ^* and the corresponding pair of optimal gray scale values $\mu^* = (\mu^*_{\text{int}}, \mu^*_{\text{ext}})$ are defined as the solution of the variational problem

$$E_{CV}(\mu^*, \Gamma^*) = \min_{\mu,\Gamma} E_{CV}(\mu, \Gamma), \tag{2.1}$$

where E_{CV} is the well-known Chan-Vese functional

$$\begin{aligned} E_{CV}(\mu, \Gamma) = \alpha \int_\Gamma d\sigma + \beta \Big\{ & \frac{1}{2} \int_{\text{int}(\Gamma)} (I(x) - \mu_{\text{int}})^2 \, dx \\ & + \frac{1}{2} \int_{\text{ext}(\Gamma)} (I(x) - \mu_{\text{ext}})^2 \, dx \Big\}, \end{aligned} \tag{2.2}$$

which corresponds to the piecewise constant Mumford-Shah model [106] in the special case of two subregions. Here $d\sigma$ denotes the arc length element, and $\alpha, \beta > 0$ are weight parameters. The first term in E_{CV} is the total length of the contour: It serves to regularize the optimal contour. The second term is the fidelity term, which penalizes deviations of the piecewise constant image model from the actual image I.

For any fixed contour Γ, not necessarily the optimal one, it turns out that the best choice of the gray scale values $\mu = (\mu_{\text{int}}, \mu_{\text{ext}})$ corresponds to the mean value of the pixel values inside and the outside Γ, respectively

$$\mu_{\text{int}} = \mu_{\text{int}}(\Gamma) = \frac{1}{|\text{int}(\Gamma)|} \int_{\text{int}(\Gamma)} I(x) \, dx, \tag{2.3}$$

$$\mu_{\text{ext}} = \mu_{\text{ext}}(\Gamma) = \frac{1}{|\text{ext}(\Gamma)|} \int_{\text{ext}(\Gamma)} I(x) \, dx. \tag{2.4}$$

Here the symbol $|A|$ denotes the area of the subset $A \subset \mathbb{R}^2$. Now, if we introduce the so-called

"reduced" Chan-Vese functional

$$E_{CV}^{R}(\Gamma) := E_{CV}(\mu(\Gamma),\Gamma), \tag{2.5}$$

then the optimal contour Γ^* can be found by solving the simpler minimization problem

$$E_{CV}^{R}(\Gamma^*) = \min_{\Gamma} E_{CV}^{R}(\Gamma). \tag{2.6}$$

Once Γ^* is found we have $\mu^* = \mu(\Gamma^*)$, of course. The minimization problem in (2.6) is solved using a gradient descent procedure, which will be recalled in the Subsection 2.2.3, after the material on the level set representation and the kinematics of moving surfaces have been presented.

2.2.2 The Level Set Method

A simple closed curve Γ can be represented as the zero level set of a function $\phi : \mathbb{R}^2 \to \mathbb{R}$ as

$$\Gamma = \{x \in \mathbb{R}^2 \,;\, \phi(x) = 0\}. \tag{2.7}$$

The sets $\mathrm{int}(\Gamma) = \{x \,;\, \phi(x) < 0\}$ and $\mathrm{ext}(\Gamma) = \{x \,;\, \phi(x) \geq 0\}$ are then the inside and the outside of Γ, respectively. Geometric quantities such as the outward unit normal $\mathbf{n}$ and the curvature κ can be expressed in terms of ϕ as

$$\mathbf{n} = \frac{\nabla\phi}{|\nabla\phi|} \quad \text{and} \quad \kappa = \nabla \cdot \frac{\nabla\phi}{|\nabla\phi|}. \tag{2.8}$$

The function ϕ is usually called the *level set function* for Γ, cf. e.g. [110].

A curve evolution, that is, a time dependent curve $t \mapsto \Gamma(t)$ can be represented by a time dependent level set function $\phi : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ as $\Gamma(t) = \{x \in \mathbb{R}^2 \,;\, \phi(x,t) = 0\}$. Let us consider the kinematics of curve evolutions. It does not make sense to "track" points as there is no way of knowing the tangential motion of points on $\Gamma(t)$. The important notion is that of *normal velocity*. The normal velocity of a curve evolution $t \mapsto \Gamma(t)$ is the scalar function defined by

$$v(\Gamma)(x) = \frac{d}{dt}\Gamma(t)(x) := -\frac{\partial\phi(x,t)/\partial t}{|\nabla\phi(x,t)|} \qquad (x \in \Gamma(t)). \tag{2.9}$$

The normal velocity is independent of the curve representation, in particular of the choice of level set function ϕ for Γ, and is therefore a geometric property of the evolution, cf. [138].

The set of possible normal velocities $v = v(\Gamma)$ of moving contours $t \mapsto \Gamma(t)$ passing through the contour Γ at time $t = 0$ is an infinite dimensional vector space. This vector space can be endowed with a natural scalar product and a corresponding norm, cf. [138]

$$\langle v, w \rangle_\Gamma = \int_\Gamma v(x) w(x)\, d\sigma \quad \text{and} \quad \|v\|_\Gamma^2 = \langle v, v \rangle_\Gamma, \tag{2.10}$$

where v, w are normal velocities. In the following we therefore denote the vector space of normal velocities at Γ by $L^2(\Gamma)$.

2.2.3 Gradient Descent Evolution

The scalar product (2.10) is important in the construction of gradient descent flows for energy functionals $E(\Gamma)$ defined on curves. Suppose $v \in L^2(\Gamma)$ is a fixed normal velocity, and let $t \mapsto \Gamma(t)$ be any moving contour which satisfies $\Gamma(0) = \Gamma$, and $(d/dt)\Gamma(0) = v$. Then the Gâteaux variation $dE(\Gamma)v$ of the functional $E = E(\Gamma)$ at the contour Γ is defined as the derivative

$$dE(\Gamma)v := \frac{d}{dt} E(\Gamma(t))\Big|_{t=0}. \tag{2.11}$$

Suppose there exists a function $\nabla E(\Gamma) \in L^2(\Gamma)$ such that E's Gâteaux variation $dE(\Gamma)v$ at Γ can be expressed in terms of the scalar product (2.10) in the following manner

$$dE(\Gamma)v = \langle \nabla E(\Gamma), v \rangle_\Gamma \quad \text{for all } v \in L^2(\Gamma). \tag{2.12}$$

Then the vector $\nabla E(\Gamma)$ is called the L^2-*gradient* of E at Γ. It is unique if it exists. The gradient descent flow for the problem of minimizing $E(\Gamma)$ can now be defined as the initial value problem

$$\frac{d}{dt}\Gamma(t) = -\nabla E(\Gamma(t)), \qquad \Gamma(0) = \Gamma_0, \tag{2.13}$$

where Γ_0 is an initial contour specified by the user.

As an example, relevant for the application in this work, notice that the L^2-gradient of the reduced functional E_{CV}^R defined in (2.5) is given by

$$\nabla E_{CV}^R(\Gamma) = \alpha\kappa + \beta\big[\frac{1}{2}(I - \mu_{\text{int}}(\Gamma))^2 - \frac{1}{2}(I - \mu_{\text{ext}}(\Gamma))^2\big], \tag{2.14}$$

where $\kappa = \kappa(x)$ is the curvature at $x \in \Gamma$. If we combine the definition of gradient descent evolution in (2.13) with (2.9) for the normal velocity, then we get the gradient descent procedure

in the level set framework

$$\frac{\partial \phi}{\partial t} = \Big(\alpha\kappa + \beta\big[\frac{1}{2}(I-\mu_{\text{int}}(\Gamma))^2 - \frac{1}{2}(I-\mu_{\text{ext}}(\Gamma))^2\big]\Big)|\nabla\phi|, \tag{2.15}$$

with $\phi(x,0) = \phi_0(x)$, where ϕ_0 is the level set function for the initial contour Γ_0. It is understood that the gray scale values $\mu_{\text{int}}(\Gamma)$ and $\mu_{\text{ext}}(\Gamma)$ are given by (2.3) and (2.4), respectively.

2.3 Remarks about Minimization of the Chan-Vese Functional

As stated in the previous section, the segmentation of a given image is obtained by finding the pair (Γ^*, μ^*) which solves the minimization problem of (2.1), where $\mu \in \mathbb{R}^2$ and Γ range over the set of admissible contours. The commonest and most straight forward way to obtain the solution of (2.1) is to apply a gradient descent procedure to the functional E_{CV}. In our particular case, two different gradient descent procedures can be constructed: In the first of these, the functional $E_{CV} = E_{CV}(\Gamma, \mu)$ is minimized simultaneously with respect to both the contour Γ and the intensity parameters μ, by finding a path $t \mapsto (\Gamma(t), \mu(t))$ which solves the system of descent differential equations

$$\begin{cases} \dfrac{d}{dt}\mu(t) = -\nabla_\mu E_{CV}(\Gamma(t), \mu(t)) \\ \dfrac{d}{dt}\Gamma(t) = -\nabla_\Gamma E_{CV}(\Gamma(t), \mu(t)) \end{cases}, \tag{2.16}$$

where the initial state $(\Gamma(0), \mu(0)) = (\Gamma_0, \mu_0)$ is specified by the user.

The second gradient descent method, which is actually the one described in the previous section, is based on the following observation: If we choose an admissible contour Γ and keep it fixed, then the fidelity term of the Chan-Vese functional

$$\frac{1}{2}\int_{\text{int}(\Gamma)} (I(x)-\mu_{\text{int}})^2\,dx + \frac{1}{2}\int_{\text{ext}(\Gamma)} (I(x)-\mu_{\text{ext}})^2\,dx, \tag{2.17}$$

is a simple quadratic function of $\mu = (\mu_{\text{int}}, \mu_{\text{ext}}) \in \mathbb{R}^2$. Therefore $\mu \mapsto E_{CV}(\Gamma, \mu)$ is minimized (uniquely) by the vector $\mu = \mu(\Gamma)$ whose components are given by (2.3) and (2.4). Using this result we find that

$$\begin{aligned} \min_{\Gamma,\mu} E_{CV}(\Gamma,\mu) &= \min_{\Gamma}\Big\{\min_{\mu} E_{CV}(\Gamma,\mu)\Big\} \\ &= \min_{\Gamma} E_{CV}(\Gamma, \mu(\Gamma)). \end{aligned} \tag{2.18}$$

Therefore, if we define the so-called "reduced" Chan-Vese functional by (2.5), then we see that the segmentation is given by the contour Γ^* which solves the minimization problem of (2.6).

In order to implement the gradient descent scheme (2.13) for the reduced Chan-Vese functional we must compute the L^2-gradient of $E^R_{CV}(\Gamma)$. To do this we apply the chain rule to the definition of the reduced functional

$$\begin{aligned}\nabla_\Gamma E^R_{CV}(\Gamma) &= \nabla_\Gamma E_{CV}(\Gamma, \mu(\Gamma)) \\ &= (\nabla_\Gamma E_{CV})(\Gamma, \mu(\Gamma)) + (\nabla_\mu E_{CV})(\Gamma, \mu(\Gamma)) \cdot (\nabla_\Gamma \mu)(\Gamma),\end{aligned} \tag{2.19}$$

where $(\nabla_\Gamma \mu)(\Gamma)$ denotes the L^2-gradient of the vector-valued functional $\Gamma \mapsto \mu(\Gamma) \in \mathbb{R}^2$. It seems as if we have to compute $(\nabla_\Gamma \mu)(\Gamma)$ in order to get the L^2-gradient of the reduced functional. It is important, however, to recall that $\mu(\Gamma)$ minimizes $\mu \mapsto E_{CV}$, hence

$$(\nabla_\mu E_{CV})(\Gamma, \mu(\Gamma)) = 0, \tag{2.20}$$

hence we have proven the often overlooked fact that

$$\nabla_\Gamma E^R_{CV}(\Gamma) = (\nabla_\Gamma E_{CV})(\Gamma, \mu(\Gamma)), \tag{2.21}$$

which explains the gradient descent procedure introduced in the previous section.

2.4 Segmentation of Image Sequences

2.4.1 A Variational Updating-Model

In this section we are going to present the basic principles behind our variational model for updating segmentation results from one frame to the next in an image sequence. See [67] for details.

Let $I_j : \Omega \to \mathbb{R}$, $j = 1, \ldots, N$, be a succession of frames from a given image sequence. Also, for some integer k, $1 \leq k \leq N$, suppose that all the frames $I_1, I_2, \ldots, I_{k-1}$ have already been segmented, such that the corresponding contours $\Gamma_1, \Gamma_2, \ldots, \Gamma_{k-1}$ are available. In order to take advantage of the prior knowledge obtained from earlier frames in the segmentation of I_k, we propose the following method: If $k = 1$, i.e. if no previous frames have actually been segmented, then we just use the classical Chan-Vese model, as presented in Section 2.2. If

$k > 1$, then the segmentation of I_k is given by the contour Γ_k which minimizes an *augmented* Chan-Vese functional of the form

$$E^A_{CV}(\Gamma_{k-1},\Gamma) := E^R_{CV}(\Gamma) + \gamma E_I(\Gamma_{k-1},\Gamma), \tag{2.22}$$

where E^R_{CV} is the reduced Chan-Vese functional defined in (2.5), $E_I = E_I(\Gamma_{k-1},\Gamma)$ is an *interaction term*, which penalizes deviations of the current active contour Γ from the previous one, Γ_{k-1}, and $\gamma > 0$ is a coupling constant which determines the strength of the interaction. The precise definition of E_I is described in the next subsection.

Figure 2.1: The initialization (green contour) and segmentation results of a person (red contour) in an image sequence using the Chan-Vese model.

2.4.2 The Interaction Term

The interaction $E_I(\Gamma_0,\Gamma)$ between a fixed contour Γ_0 and an active contour Γ, used in (2.22), may be chosen in several different ways. Two common choices are the so-called pseudo-distances, cf. [42], and the *area of the symmetric difference* of the sets $\text{int}(\Gamma)$ and $\text{int}(\Gamma_0)$, cf. [32]. We have found that none of the mentioned contour interactions satisfy our needs, and we have therefore chosen to introduce a new pose-invariant interaction term.

To describe this interaction term, let $\phi_0 : \Omega \to \mathbb{R}$ denote the *signed distance function* associated with the contour Γ_0, that is, the function

$$\phi_0(x) = \begin{cases} \text{dist}(x,\Gamma_0) & \text{for } x \in \text{ext}(\Gamma_0), \\ -\text{dist}(x,\Gamma_0) & \text{for } x \in \text{int}(\Gamma_0). \end{cases} \tag{2.23}$$

Then the interaction $E_I = E_I(\Gamma_0;\Gamma)$ is defined by the formula

$$E_I(\Gamma_0,\Gamma) = \min_T \int_{\text{int}(\Gamma)} \phi_0(T^{-1}x)\,dx, \tag{2.24}$$

where the minimum is taken over the *group of Euclidean transformations* $T : \mathbb{R}^2 \to \mathbb{R}^2$ which preserves the orientation of the plane, that is, transformations T which are compositions of translations and rotations (but not reflections). Minimizing over groups of transformations is a standard devise to obtain pose-invariant interactions, see [32] and [42].

For any given contour Γ, let $T = T(\Gamma)$ denote the transformation which minimizes the expression on the right hand side of (2.24). Since this is an optimization problem $T(\Gamma)$ can be found using gradient descent. For simplicity of presentation, suppose we only consider the group of translations $T_{\mathbf{a}} : x \mapsto x + \mathbf{a}$, $\mathbf{a} \in \mathbb{R}^2$, and want to determine the optimal translation vector $\mathbf{a} = \mathbf{a}(\Gamma)$. Then we have to solve the optimization problem

$$\min_{\mathbf{a}} \int_{\text{int}(\Gamma)} \phi_0(x - \mathbf{a})\, dx.$$

The optimal translation $\mathbf{a}(\Gamma)$ can then be obtained as the limit, as time t tends to infinity, of the solution to initial value problem

$$\dot{\mathbf{a}}(t) = \int_{\text{int}(\Gamma)} \nabla\phi_0(x - \mathbf{a}(t))\, dx \ , \qquad \mathbf{a}(0) = 0. \tag{2.25}$$

Similar gradient descent schemes can be devised for rotations and scalings (in the case of similarity transformation), cf. [32].

2.4.3 The Gradient Descent Equations

The augmented Chan-Vese functional (2.22) is minimized using standard gradient descent as described in Section 2.2. That is, we solve the initial value problem

$$\begin{aligned} \frac{d}{dt}\Gamma(t) &= -\nabla E^A_{CV}(\Gamma_{k-1}, \Gamma(t)) \\ &:= -\nabla E^R_{CV}(\Gamma_{k-1}, \Gamma(t)) - \gamma \nabla E_I(\Gamma_{k-1}; \Gamma(t)), \end{aligned} \tag{2.26}$$

with the initial contour $\Gamma(0) = \Gamma_{k-1}$, and pass to the limit $t \to \infty$. Here ∇E^R_{CV} is the L^2-gradient of the reduced Chan-Vese functional, see Eq. (2.14), and ∇E_I is the L^2-gradient of the interaction E_I, which is given by the formula

$$\nabla E_I(\Gamma_{k-1}, \Gamma; x) = \phi_{k-1}(T(\Gamma)x), \qquad (\text{for } x \in \Gamma), \tag{2.27}$$

as is easily verified. Here ϕ_{k-1} is the signed distance function for Γ_{k-1}.

Notice that the L^2-gradient of the reduced functional E^R_{CV} defined in (2.5) is given by (2.14). If we combine the definition of gradient descent evolutions in (2.13) with (2.9) for the normal velocity, then we get the gradient descent procedure for the proposed model in the level set framework

$$\begin{aligned}\frac{\partial \phi}{\partial t} &= \Big(\alpha\kappa+\beta\big[\frac{1}{2}(I-\mu_{\text{int}}(\Gamma))^2-\frac{1}{2}(I-\mu_{\text{ext}}(\Gamma))^2\big] \\ &\quad +\gamma\phi_{k-1}(T(\Gamma)x)\Big)|\nabla\phi|,\end{aligned} \tag{2.28}$$

with $\phi(x,0)=\phi_0(x)$, where ϕ_0 is the level set function for the initial contour Γ_0. Again it is understood that the gray scale values $\mu_{\text{int}}(\Gamma)$ and $\mu_{\text{ext}}(\Gamma)$ are given by (2.3) and (2.4), respectively.

2.5 Remarks about the Interaction Term

In this section we give a motivation for the particular choice of shape dissimilarity measure that we use for the contour interaction term in this work.

Let us recall the setting: As before, let Γ denote the active contour and Γ_0 the contour corresponding to the shape prior. Also, let $\phi=\phi(x)$ and $\phi_0=\phi_0(x)$ denote the signed distance functions associated with Γ and Γ_0, respectively, where x is a generic point in the image domain Ω. Notice that, for any $x\in\Omega$ fixed, $\phi(x)=\phi(\Gamma;x)$ is actually a functional of the contour Γ. This observation will be important in the discussion below.

Now, if we disregard pose invariance for the moment, (e.g. by assuming that Γ_0 is already optimally aligned with Γ in the appropriate sense,) then the shape interaction term proposed in this work has the form

$$E_{\text{shape}}(\Gamma,\Gamma_0)=\int_{\text{int}(\Gamma)}\phi_0(x)\,dx. \tag{2.29}$$

We wish to compare (2.29) with two other popular choices for the shape dissimilarity measure. The first of these is the classic *area of the symmetric difference*, which has been used in [32], and [126]

$$E^{\text{SD}}_{\text{shape}}(\Gamma,\Gamma_0)=\text{area}(\Sigma\triangle\Sigma_0). \tag{2.30}$$

Here we have used the notation $\Sigma\triangle\Sigma_0:=(\Sigma\cup\Sigma_0)\backslash(\Sigma\cap\Sigma_0)$ to denote the symmetric difference

of the two sets $\Sigma = \text{int}(\Gamma)$, $\Sigma_0 = \text{int}(\Gamma_0)$. The second one is the so-called pseudo-distance

$$E^{\text{PD}}_{\text{shape}}(\Gamma,\Gamma_0) = \frac{1}{2}\int_\Omega [\phi(x) - \phi_0(x)]^2\,dx, \tag{2.31}$$

which has been studied, with various minor modifications, in [128], [117], and [42].

2.5.1 Area of the Symmetric Difference

As a dissimilarity measure, *area of the symmetric difference* (2.30) has several desirable features: it is symmetric with respect to the contours Γ and Γ_0, that is, $E^{\text{SD}}_{\text{shape}}(\Gamma,\Gamma_0) = E^{\text{SD}}_{\text{shape}}(\Gamma_0,\Gamma)$, and $E^{\text{SD}}_{\text{shape}}(\Gamma,\Gamma_0) = 0$ if and only if $\Gamma = \Gamma_0$. Moreover, for Γ_0 fixed, it is easy to compute the L^2-gradient of $E^{\text{SD}}_{\text{shape}}$ with respect to Γ. In fact, if $\mathbf{1}_\Sigma$ and $\mathbf{1}_{\Sigma_0}$ denote the characteristic functions (or indicator functions) of Σ and Σ_0, respectively, then we may write (2.30) as

$$\begin{aligned} E^{\text{SD}}_{\text{shape}}(\Gamma,\Gamma_0) &= \int_\Omega [\mathbf{1}_\Sigma(x) - \mathbf{1}_{\Sigma_0}(x)]^2\,dx \\ &= \int_\Omega [\mathbf{1}_\Sigma(x) - 2\mathbf{1}_\Sigma(x)\mathbf{1}_{\Sigma_0}(x) + \mathbf{1}_{\Sigma_0}(x)^2]\,dx \\ &= \int_\Omega [1 - 2\mathbf{1}_{\Sigma_0}(x)]\,dx + \text{const.}, \end{aligned}$$

so $E^{\text{SD}}_{\text{shape}}$ is an area-based functional whose Gâteaux derivative is

$$dE^{\text{SD}}_{\text{shape}}(\Gamma,\Gamma_0)v = \int_\Gamma [1 - 2\cdot\mathbf{1}_{\Sigma_0}(x)]v(x)\,d\sigma(x), \tag{2.32}$$

for any normal variation $v : \Gamma \to \mathbb{R}$ of the active contour. It follows that the L^2-gradient is given by

$$\nabla_\Gamma E^{\text{SD}}_{\text{shape}}(\Gamma,\Gamma_0) = 1 - 2\mathbf{1}_{\Sigma_0}(x) \quad (x \in \Gamma). \tag{2.33}$$

This formula reveals a shortcoming of $E^{\text{SD}}_{\text{shape}}$ as interaction term; The L^2-gradient has the same magnitude, ± 1, no matter if Γ is close to the prior shape Γ_0 or not. So if $E^{\text{SD}}_{\text{shape}}$ is used as dissimilarity measure in a segmentation model with priors, then the prior will either be ignored, if the coupling constant has a value below a certain threshold value, or the prior will overrule the shape information contained in the fidelity term, if the coupling constant is greater than the threshold value. This means that $E^{\text{SD}}_{\text{shape}}$ is unsuitable for segmentation models where prior information should be adaptable to image information.

Notice that the L^2-regularization term, proposed for image registration in a paper by Lukas

and Kanade [96], is exactly equal to $E^{\mathrm{SD}}_{\mathrm{shape}}(\Gamma,\Gamma_0)$ when shapes are represented by binary images, as above.

2.5.2 The Pseudo-Distance

The pseudo-distance (2.31) is symmetric, i.e., $E^{\mathrm{PD}}_{\mathrm{shape}}(\Gamma,\Gamma_0) = E^{\mathrm{PD}}_{\mathrm{shape}}(\Gamma_0,\Gamma)$, and equals zero if and only if the contours are equal. The main drawback with the pseudo-distance is that it is difficult to compute its L^2-gradient, which is needed in gradient descent PDEs for functionals containing the pseudo-distance. To explain this problem, assume that Γ_0 is fixed, let the function $v : \Gamma \to \mathbb{R}$ be a normal variation of the active contour Γ, and observe that the Gâteaux derivative of $E^{\mathrm{PD}}_{\mathrm{shape}}$ with respect to Γ is given by

$$dE^{\mathrm{PD}}_{\mathrm{shape}}(\Gamma,\Gamma_0)v = \int_\Omega (\phi(\Gamma,x) - \phi_0(x)) d\phi(\Gamma;x)v\,dx, \tag{2.34}$$

where $d\phi(\Gamma;x)v$ denotes the Gâteaux derivative of the functional $\Gamma \to \phi(\Gamma;x)$, $x \in \Omega$ fixed, which must be computed. For $x \in \Omega$ let $P_\Gamma x$ denote the point on Γ which is closest to x. The mapping $\Omega \ni x \to P_\Gamma x \in \Gamma$ is well-defined for almost all points in Ω, and we refer to $P_\Gamma x$ as the projection of x onto Γ. The Gâteaux derivative with respect to Γ of the signed distance mapping can be expressed in terms of the projection mapping as

$$d\phi(\Gamma;x)v = -v(P_\Gamma x), \tag{2.35}$$

so (2.34) becomes

$$dE^{\mathrm{PD}}_{\mathrm{shape}}(\Gamma,\Gamma_0)v = -\int_\Omega (\phi(\Gamma,x) - \phi_0(x)) v(P_\Gamma x)\,dx, \tag{2.36}$$

which is not easily expressed as a curve integral of the form $\int_\Gamma F(x)v(x)\,d\sigma(x)$. This explains the difficulty in computing $\nabla_\Gamma E^{\mathrm{PD}}_{\mathrm{shape}}$. The derivatives computed in [42, 117, 128] are not Gâteaux derivatives (or shape derivatives) in our sense, c.f. [138] or [50], but rather the first variation of $E^{\mathrm{PD}}_{\mathrm{shape}}$ with respect to ϕ, computed as if the set of signed distance functions ϕ is a vector space.

2.5.3 Proposed Interaction Term

The main benefit of our interaction term E_{shape} defined in (2.29) is that its L^2-gradient can be computed easily by

$$\nabla_\Gamma E_{\text{shape}}(\Gamma,\Gamma_0) = \phi_0(x) = \phi(\Gamma_0;x) \quad (x \in \Gamma),$$

and that this gradient is small if Γ is close to the shape prior Γ_0, and large if the active contour is far from the shape prior. However, $E_{\text{shape}}(\Gamma,\Gamma_0)$ is not symmetric in Γ and Γ_0, which may in general be considered a drawback. However, in our particular application, where we want to use shape information from a previous image frame (Γ_0) to guide the segmentation in the current frame (Γ), the lack of symmetry does not seem to be such a big issue. After all, there is no obvious symmetry between *past* and *present!* Suppose instead that we wanted to segment an image sequence simultaneously, by considering the stack of frames as a three-dimensional object, then it would be relevant to use a symmetric interaction term between the contours in the individual frames.

2.6 Experimental Results

In this section we present the results obtained from experiments using synthetic and real image sequences. We use the Chan-Vese model to segment a selected object with approximately uniform intensity of the first frame of each image sequence. The proposed method is then applied to segment the next frames of the image sequences sequentially frame-by-frame, where the segmentation result in one frame is used as the initial contour in the next one. Since we do not have the prior model of the object of interest, it is important to get a good result of the first frame in order to get better results in the next frames. The minimization of the functional, giving the optimal contour, is obtained from the gradient descent procedure (2.26) which has been implemented in the level set framework outlined in Section 2.2. In all experiments, the initial contour in the first frame is given manually, as a circle, located inside the object of interests. Note that not all frames of each image sequences are shown in the figures.

As illustrated in Figure 2.1, the Chan-Vese model is capable of segmenting an object of interest in an image sequence without any problems. Further such results can be found in [103]. However, as pointed out in the reference, the Chan-Vese method will have problems if occlusions appear in the image which cover the whole or parts of the object being segmented.

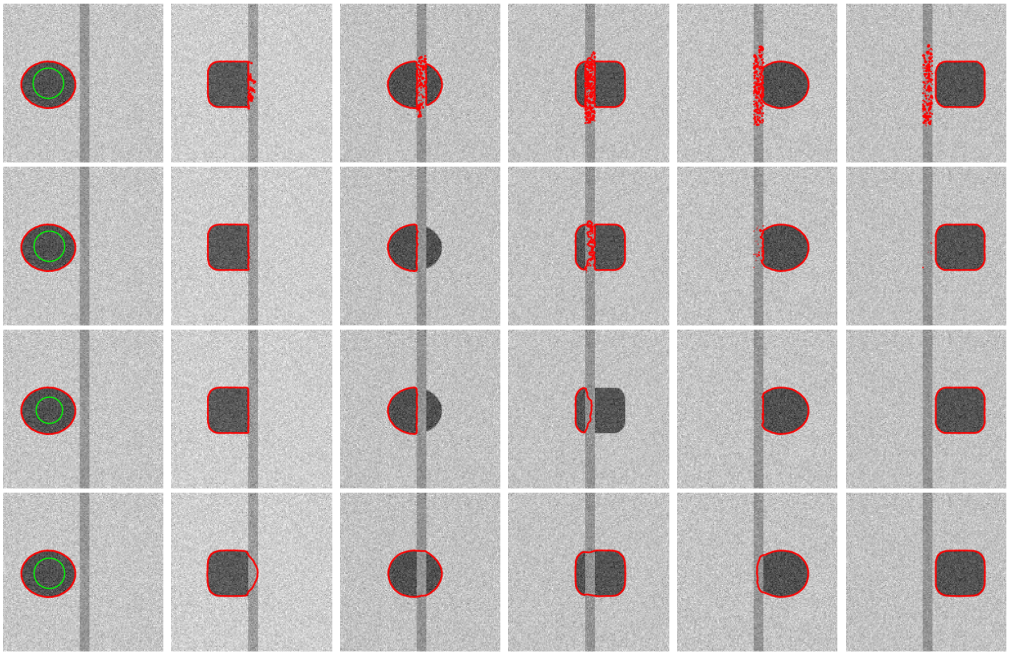


Figure 2.2: The initialization (green contour) and segmentation results (red contour) of a nonrigid object in a synthetic image sequence with additive Gaussian noise. First, second, and third rows: without interaction term and increasing α. Fourth row: with interaction term.

Figure 2.2 shows the segmentation results for a nonrigid object in a synthetic image sequence, where occlusions appear. The Chan-Vese method fails to segment the object of interest when it reaches the occlusion, even when we increase the regularization weight α to get more regular results (Rows 1-3). Using the proposed method, we obtain much better results (Last Row).

In Figure 2.3, we show the experiment of segmenting a car (white van) which passes through the obstacle in a traffic sequence. The Chan-Vese method fails to segment the car when it reaches the lamp post (first column). This result can then be improved by increasing α, but it is still unable to segment the car when partly occluded, instead the contour is disappearing near the lamp post when α is too large (second column). This problem can be overcome by adding the interaction term (third and fourth column), where we use different value of α to see the influence of the regularization term coupled with the interaction term with the same value of γ. As we can see from Fig. 2.4 using large α the contour is unable to pass the lamp post while we get better results by using the proposed method, even with small α.

Another experiment is given in Figure 2.5 where a walking person being segmented is partly covered by another person. Without interaction term, the segmentation results are poor when we use small α (first column). The results can then be improved using larger α, but

Algorithm 1 The algorithm for segmentation of N frames image sequence from the second frame $I_2 ... I_N$.

INPUT: Current frame I_k and the level set function from the previous frame ϕ_{k-1}
OUTPUT: Optimal level set function ϕ_k.

1. **Initialization** Initialize the level set function $\phi_k = \phi_{k-1}$.
2. **Computation** Compute the optimal translation vector and then the gradient descent of (2.26).
3. **Re-initialization** Re-initialize the level set function ϕ_k.
4. **Convergence** Stop if the level set evolution converges, otherwise go to step 2.

this limits the deformation of the contour especially in the first frame (second column). The proposed method prevents the segmentation of the wrong object and give better results, as is clearly shown. In this experiment, the coupling constant γ is varied to see the influence of the interaction term on the results (third and fourth columns). The contour is only slightly affected by the prior if γ is small. On the other hand, if γ is too large, the contour will be close to a similarity transformed version of the prior.

2.7 Conclusions

We have presented a novel method for variational segmentation of image sequences containing nonrigid, moving objects. The proposed method is formulated as variational problem, with one part of the functional corresponding to the Chan-Vese model and another part corresponding to the pose-invariant interaction with a shape prior based on the previous contour. The optimal transformation as well as the shape deformation are determined by minimization of an energy functional using a gradient descent scheme. Experimental results are shown on synthetic and real image sequences and its performance looks promising.

Acknowledgements

The human walking sequences were downloaded from EU funded CAVIAR project (IST 2001 37540) website and the traffic sequence from KOGS/IAKS Universität Karlsruhe.

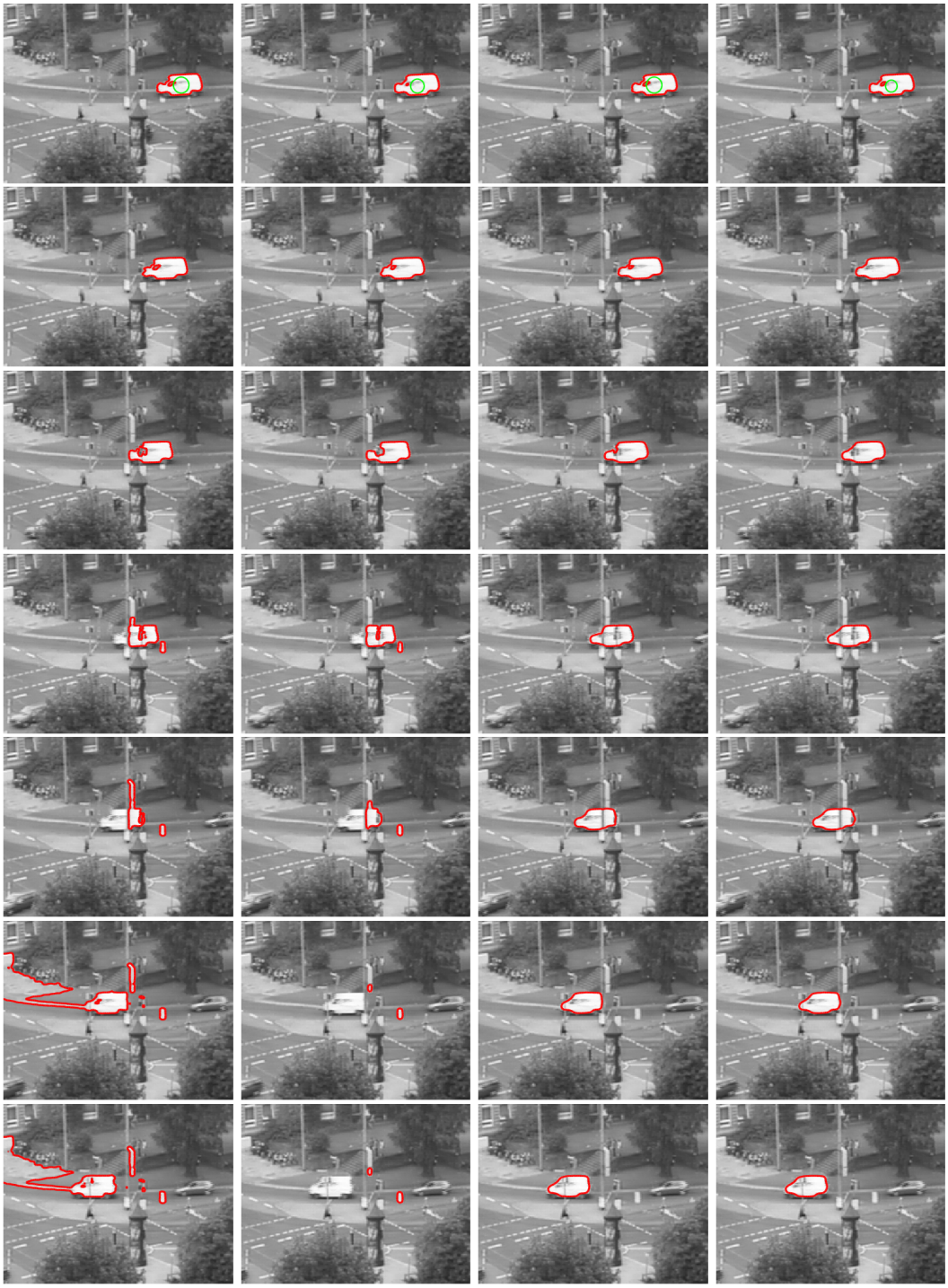


Figure 2.3: The initialization (green contour) and segmentation results (red contour) of a car which passes through a lamp post in a traffic sequence. First and second columns: without interaction term and increasing α. Third and fourth columns: with interaction term and increasing α.

Figure 2.4: Segmentation results taken from the fourth row of Figure 2.3. Left: without interaction term. Right: with interaction term.

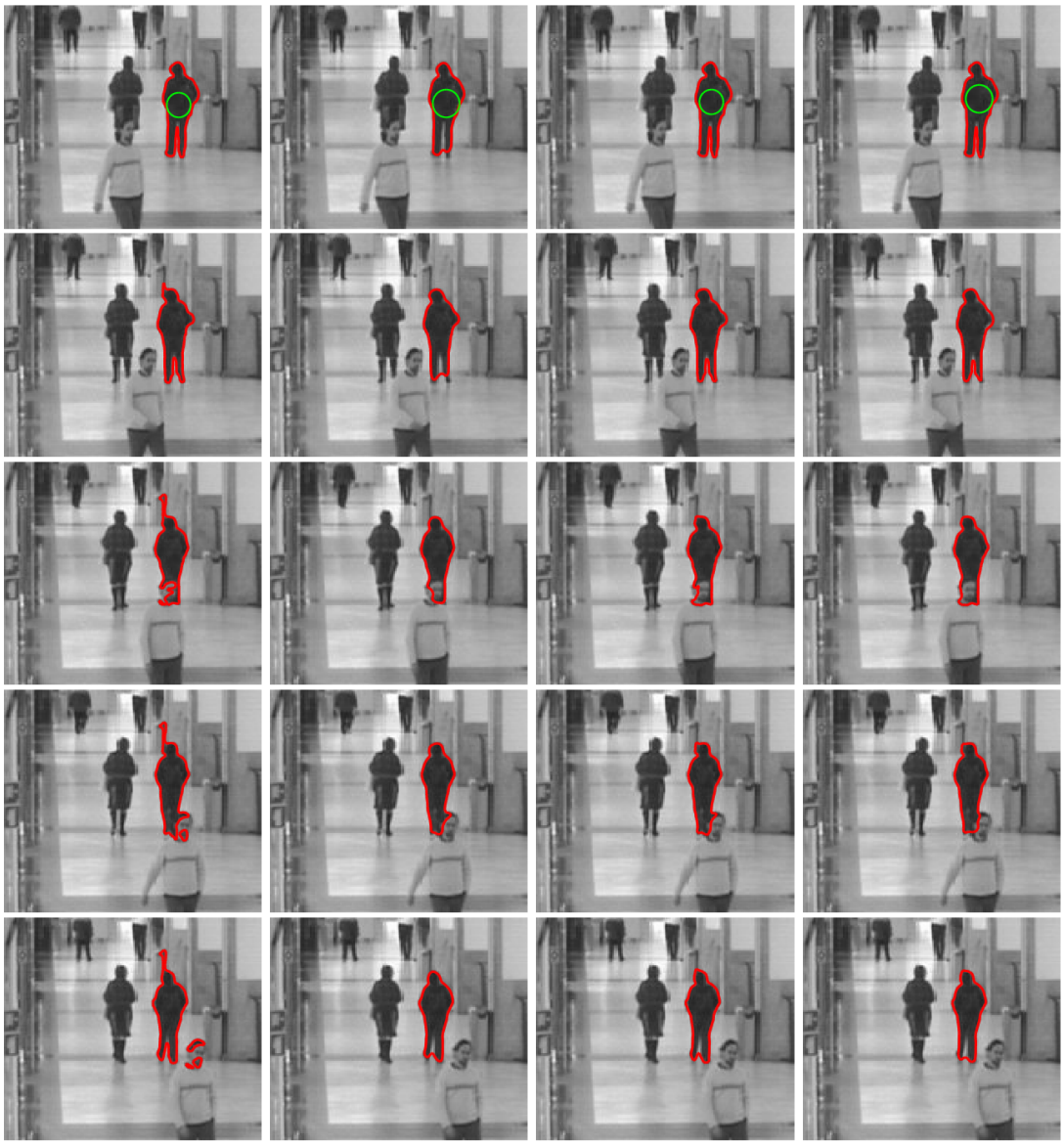


Figure 2.5: The initialization (green contour) and segmentation results (red contour) of a person that partly occluded by another person in a human walking sequence. First and second columns: without interaction term and increasing α. Third and fourth columns: with interaction term and increasing γ.

Chapter 3

Nonrigid Variational Object Segmentation with Occlusion Detection in Image Sequences

---

This chapter is the result of merging papers [68, 74] with minor modifications without affecting the content or the results. Here we address the problem of nonrigid object segmentation in image sequences in the presence of occlusions. The proposed variational segmentation method is based on a region-based active contour of the Chan-Vese model augmented with a frame-to-frame interaction term as a shape prior. The interaction term is constructed to be pose-invariant by minimizing over a group of transformations and to allow moderate deformation in the shape of the contour. The segmentation method is then coupled with a novel variational contour matching formulation between two consecutive contours which gives a mapping of the intensities from the interior of the previous contour to the next. With this information occlusions can be detected and located using deviations from predicted intensities and the missing intensities in the occluded regions can be reconstructed. After reconstructing the occluded regions in the novel image, the segmentation can then be improved. Experimental results on synthetic and real image sequences are shown.

3.1 Introduction

Object segmentation is one of the most important processes in computer vision which aims at extracting the object of interests lying in the image. This is a very difficult process since the object of interests could be diverse, complex and the understanding on them vary according to each individual. The process becomes more difficult when the objects to be segmented are moving and nonrigid and even more when occlusions appear. The shape of nonrigid, moving objects may vary a lot along image sequences due to, for instance, deformations or occlusions, which puts additional constraints on the segmentation process.

Numerous methods have been proposed and applied to this problem. Active contours are powerful methods for image segmentation; either boundary-based such as geodesic active contours [27], or region-based such as Chan-Vese models [31], which are formulated as variational problems. Those variational formulations perform quite well and have often been applied based on level sets. Active contour based segmentation methods often fail due to noise, clutter and occlusion. In order to make the segmentation process robust against these effects, shape priors have been proposed to be incorporated into the segmentation process, such as in [22, 32, 35, 39, 42, 43, 93, 128, 146]. However, major occlusions is still a big problem. In order to improve the robustness of the segmentation methods in the presence of occlusions, it is necessary to detect and locate the occlusions [70, 88, 142]. Then using this information, the segmentation can be improved. For example, [144] proposed that the spatial order information in the image model is used to impose dynamically shape prior constraints only to occluded boundaries.

This chapter focuses on the region-based variational approach to segment a nonrigid object that may be partially occluded in image sequences. In particular, we would like to distinguish real shape deformations of the object from apparent shape deformations due to occlusions. We propose and analyze a novel variational segmentation method for image sequences, that can both deal with shape deformations and at the same time is robust to noise, clutter and occlusions. The proposed method is based on minimizing an energy functional containing the standard Chan-Vese functional as one part and a term that penalizes the deviation from the previous shape as a second part. The second part of the functional is based on a transformed distance map to the previous contour, where different transformation groups, such as Euclidean, similarity or affine, can be used depending on the particular application. This variational framework is then augmented with a novel contour flow algorithm, giving a mapping of the intensities inside

the contour of one image to the inside of the contour in the next image. Using this mapping, occlusions can be detected and located by simply thresholding the difference between the transformed intensities and the observed ones in the novel image. By using occlusions information, the occluded regions are reconstructed to improve the segmentation results.

3.2 Segmentation of Image Sequences

In this section, we describe the region-based segmentation model of Chan-Vese [31] and a variational model for updating segmentation results from one frame to the next in an image sequence.

3.2.1 Region-Based Segmentation

The idea of the Chan-Vese model [31] is to find a contour Γ such that the image I is optimally approximated by a gray scale value μ_{int} on $\text{int}(\Gamma)$, the *inside* of Γ, and by another gray scale value μ_{ext} on $\text{ext}(\Gamma)$, the *outside* of Γ. The optimal contour Γ^* is defined as the solution of the variational problem,

$$E_{CV}(\Gamma^*) = \min_{\Gamma} E_{CV}(\Gamma), \tag{3.1}$$

where E_{CV} is the Chan-Vese functional,

$$E_{CV}(\mu,\Gamma) = \alpha|\Gamma| + \beta\left\{\frac{1}{2}\int_{\text{int}(\Gamma)}(I(x)-\mu_{\text{int}})^2\,dx + \frac{1}{2}\int_{\text{ext}(\Gamma)}(I(x)-\mu_{\text{ext}})^2\,dx\right\}, \tag{3.2}$$

where $|\Gamma|$ is the arc length of the contour, $\alpha,\beta > 0$ are weight parameters, and

$$\mu_{\text{int}} = \frac{1}{|\text{int}(\Gamma)|}\int_{\text{int}(\Gamma)} I(x)\,dx, \tag{3.3}$$

$$\mu_{\text{ext}} = \frac{1}{|\text{ext}(\Gamma)|}\int_{\text{ext}(\Gamma)} I(x)\,dx. \tag{3.4}$$

The gradient descent flow for the problem of minimizing a functional $E_{CV}(\Gamma)$ is the solution to initial value problem:

$$\frac{d}{dt}\Gamma(t) = -\nabla E_{CV}(\Gamma(t)), \qquad \Gamma(0) = \Gamma_0, \tag{3.5}$$

where Γ_0 is an initial contour. Here $\nabla E_{CV}(\Gamma)$ is the L^2-gradient of the energy functional

$E_{CV}(\Gamma)$, cf. e.g. [138] for definitions of these notions. Then the L^2-gradient of E_{CV} is

$$\nabla E_{CV}(\Gamma) = \alpha\kappa + \beta\big[\frac{1}{2}(I-\mu_{\text{int}}(\Gamma))^2 - \frac{1}{2}(I-\mu_{\text{ext}}(\Gamma))^2\big], \tag{3.6}$$

where κ is the curvature.

In the level set framework [110], a curve evolution, $t \mapsto \Gamma(t)$, can be represented by a time dependent level set function $\phi : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ as $\Gamma(t) = \{x \in \mathbb{R}^2 \ ; \ \phi(x,t) = 0\}$, $\phi(x) < 0$ and $\phi(x) > 0$ are the regions inside and the outside of Γ, respectively. The normal velocity of $t \mapsto \Gamma(t)$ is the scalar function $d\Gamma/dt$ defined by

$$\frac{d}{dt}\Gamma(t)(x) := -\frac{\partial\phi(x,t)/\partial t}{|\nabla\phi(x,t)|} \qquad (x \in \Gamma(t)) \ . \tag{3.7}$$

Recall that the outward unit normal $\mathbf{n}$ and the curvature κ can be expressed in terms of ϕ as $\mathbf{n} = \nabla\phi/|\nabla\phi|$ and $\kappa = \nabla \cdot \big(\nabla\phi/|\nabla\phi|\big)$.

Combined with the definition of gradient descent evolutions (3.5) and the formula for the normal velocity (3.7) this gives the gradient descent procedure in the level set framework:

$$\frac{\partial\phi}{\partial t} = \Big(\alpha\kappa + \beta\big[\frac{1}{2}(I-\mu_{\text{int}}(\Gamma))^2 - \frac{1}{2}(I-\mu_{\text{ext}}(\Gamma))^2\big]\Big)|\nabla\phi|,$$

where $\phi(x,0) = \phi_0(x)$ represents the initial contour Γ_0.

3.2.2 The Interaction Term

The interaction $E_I(\Gamma_0,\Gamma)$ between a fixed contour Γ_0 and an active contour Γ may be regarded as a shape prior and be chosen in several different ways, such as the area of the symmetric difference of the sets $\text{int}(\Gamma)$ and $\text{int}(\Gamma_0)$, cf. [32], and the pseudo-distances, cf. [42].

Let $\phi = \phi(x)$ and $\phi_0 = \phi_0(x)$ denote the signed distance functions associated with Γ and Γ_0, respectively, where x is a generic point in the image domain Ω. By assuming that Γ_0 is already optimally aligned with Γ in the appropriate sense, then the interaction term proposed here has the form:

$$E_I(\Gamma,\Gamma_0) = \int_{\text{int}(\Gamma)} \phi_0(x)\,dx \ . \tag{3.8}$$

The main benefit of our interaction term defined in (3.8) is that its L^2-gradient can be computed easily:

$$\nabla_\Gamma E_I(\Gamma,\Gamma_0) = \phi_0(x) = \phi(\Gamma_0;x) \quad (x \in \Gamma)$$

and that this gradient is small if Γ is close to the shape prior Γ_0, and large if the active contour is far from the shape prior. However, $E_I(\Gamma,\Gamma_0)$ is not symmetric in Γ and Γ_0, which may in general be considered a drawback. However, in our particular application, where we want to use shape information from a previous image frame (Γ_0) to guide the segmentation in the current frame (Γ), the lack of symmetry does not seem to be such a big issue.

The proposed interaction term is constructed to be pose-invariant and to allow moderate deformations in shape. Let $\mathbf{a} \in \mathbb{R}^2$ is a group of translations. We want to determine the optimal translation vector $\mathbf{a} = \mathbf{a}(\Gamma)$, then the interaction $E_I = E_I(\Gamma_0,\Gamma)$ is defined by the formula,

$$E_I(\Gamma_0,\Gamma) = \min_{\mathbf{a}} \int_{\text{int}(\Gamma)} \phi_0(x-\mathbf{a})\,dx. \tag{3.9}$$

Minimizing over groups of transformations is the standard device to obtain pose-invariant interactions, see [32] and [42].

Since this is an optimization problem $\mathbf{a}(\Gamma)$ can be found using the gradient descent procedure. The optimal translation $\mathbf{a}(\Gamma)$ can then be obtained as the limit, as time t tends to infinity, of the solution to initial value problem

$$\dot{\mathbf{a}}(t) = \int_{\text{int}(\Gamma)} \nabla\phi_0(x-\mathbf{a}(t))\,dx \ , \qquad \mathbf{a}(0) = 0 \ . \tag{3.10}$$

Similar gradient descent schemes can be devised for rotations and scalings (in the case of similarity transforms), cf. [32].

3.2.3 Using the Interaction Term in Segmentation of Image Sequences

Let $I_j : \Omega \to \mathbb{R}$, $j = 1,\ldots,N$, be a succession of N frames from a given image sequence. Also, for some integer k, $1 \le k \le N$, suppose that all the frames $I_1, I_2, \ldots, I_{k-1}$ have already been segmented, such that the corresponding contours $\Gamma_1,\Gamma_2,\ldots,\Gamma_{k-1}$ are available. In order to take advantage of the prior knowledge obtained from earlier frames in the segmentation of I_k, we propose the following method: If $k = 1$, i.e. if no previous frames have actually been segmented, then we just use the standard Chan-Vese model, as presented in Sect. 3.2.1. If $k > 1$, then the segmentation of I_k is given by the contour Γ_k which minimizes an *augmented* Chan-Vese functional of the form,

$$E^A_{CV}(\Gamma_{k-1},\Gamma_k) := E_{CV}(\Gamma_k) + \gamma E_I(\Gamma_{k-1},\Gamma_k), \tag{3.11}$$

where E_{CV} is the Chan-Vese functional, $E_I = E_I(\Gamma_{k-1}, \Gamma_k)$ is an *interaction term*, which penalizes deviations of the current active contour Γ_k from the previous one, Γ_{k-1}, and $\gamma > 0$ is a coupling constant which determines the strength of the interaction. See Algorithm 2.

The augmented Chan-Vese functional (3.11) is minimized using standard gradient descent (3.5) described in Sect. 3.2.1 with ∇E equal to

$$\nabla E^A_{CV}(\Gamma_{k-1}, \Gamma_k) := \nabla E_{CV}(\Gamma_k) + \gamma \nabla E_I(\Gamma_{k-1}; \Gamma_k), \tag{3.12}$$

and the initial contour $\Gamma(0) = \Gamma_{k-1}$. Here ∇E_{CV} is the L^2-gradient (3.6) of the Chan-Vese functional, and ∇E_I the L^2-gradient of the interaction term, which is given by the formula,

$$\nabla E_I(\Gamma_{k-1}, \Gamma_k; x) = \phi_{k-1}(x - \mathbf{a}(\Gamma_k)), \qquad (\text{for } x \in \Gamma_k). \tag{3.13}$$

Here ϕ_{k-1} is the signed distance function for Γ_{k-1}.

Algorithm 2 The algorithm for segmentation of N frames image sequence from the second frame $I_2 ... I_N$.

INPUT: Current frame I_k and the level set function from the previous frame ϕ_{k-1}
OUTPUT: Optimal level set function ϕ_k.

1. **Initialization** Initialize the level set function $\phi_k = \phi_{k-1}$.
2. **Computation** Compute the optimal translation vector and then the gradient descent of (3.12).
3. **Re-initialization** Re-initialize the level set function ϕ_k.
4. **Convergence** Stop if the level set evolution converges, otherwise go to step 2.

3.3 Occlusion Detection by Contour Matching

In this section we are going to present a variational solution to a contour matching problem. We start with the theory behind the contour matching problem and then describe the algorithm we use to implement it to detect and locate the occlusions.

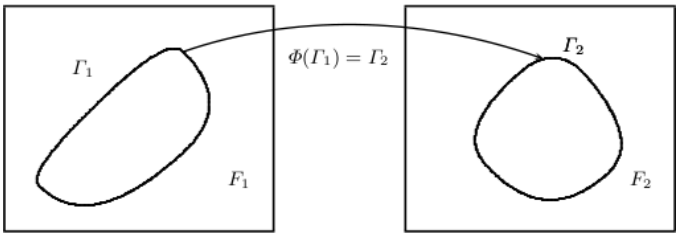


Figure 3.1: Given two closed curves Γ_1 and Γ_2 contained in two images F_1 and F_2, Φ maps F_1 onto F_2 such that Γ_1 is mapped onto Γ_2.

3.3.1 A Contour Matching Problem

Suppose we have two simple closed curves Γ_1 and Γ_2 contained in the image domain Ω. Find the "most economical" mapping $\Phi = \Phi(x) : \Omega \to \mathbb{R}^2$ such that Φ maps Γ_1 onto Γ_2, i.e. $\Phi(\Gamma_1) = \Gamma_2$, see Figure 3.1. The latter condition is to be understood in the sense that if $\alpha = \alpha(\gamma) : [0,1] \to \Omega$ is a positively oriented parametrization of Γ_1, then $\beta(\gamma) = \Phi(\alpha(\gamma)) : [0,1] \to \Omega$ is a positively oriented parametrization of Γ_2 (allowing some parts of Γ_2 to be covered multiple times).

To present our variational solution of this problem, let $\mathcal{M}$ denote the set of twice differential mappings Φ which maps Γ_1 to Γ_2 in the above sense. Loosely speaking

$$\mathcal{M} = \{\Phi \in C^2(\Omega;\mathbb{R}^2) \,|\, \Phi(\Gamma_1) = \Gamma_2\}.$$

Moreover, given a mapping $\Phi : \Omega \to \mathbb{R}^2$, not necessarily a member of $\mathcal{M}$, then we express Φ in the form $\Phi(x) = x + U(x)$, where the vector valued function $U = U(x) : \Omega \to \mathbb{R}^2$ is called the *displacement field associated with* Φ, or simply the displacement field. It is sometimes necessary to write out the components of the displacement field; $U(x) = (u_1(x), u_2(x))^T$.

We now define the "most economical" map to be the member Φ^* of $\mathcal{M}$ which minimizes the following energy functional:

$$E[\Phi] = \frac{1}{2}\int_\Omega \|DU(x)\|_F^2 \, dx \ , \tag{3.14}$$

where $\|DU(x)\|_F$ denotes the Frobenius norm of $DU(x) = [\nabla u_1(x), \nabla u_2(x)]^T$, which for an arbitrary matrix $A \in \mathbb{R}^{2\times 2}$ is defined by $\|A\|_F^2 = \mathrm{tr}(A^T A)$. That is, the optimal matching is given by

$$\Phi^* = \underset{\Phi\in\mathcal{M}}{\arg\min}\, E[\Phi] \ . \tag{3.15}$$

Using that, $E[\Phi]$ can be written in the form

$$E[\Phi] = \frac{1}{2}\int_\Omega |\nabla u_1(x)|^2 + |\nabla u_2(x)|^2\, dx \ , \tag{3.16}$$

it is easy to see that the Gâteaux derivative of $E[\Phi]$ is given by

$$\begin{aligned} dE[\Phi;V] &= \int_\Omega \nabla u_1(x)\cdot\nabla v_1(x) + \nabla u_2(x)\cdot\nabla v_2(x)\, dx \\ &= \int_\Omega \mathrm{tr}(DU(x)^T DV(x))\, dx \ , \end{aligned}$$

for any displacement field $V(x) = (v_1(x), v_2(x))^T$. After integration by parts we find that the necessary condition for $\Phi^*(x) = x + U^*(x)$ to be a solution of the minimization problem (3.15) takes the form

$$0 = -\int_\Omega \Delta U^*(x)\cdot V(x)\, dx \ , \tag{3.17}$$

for any *admissible* displacement field variation $V = V(x)$. Here $\Delta U^*(x) = (\Delta u_1(x), \Delta u_2(x))^T$ is the Laplacian of the vector valued function $U^* = U^*(x)$. Since every admissible mapping Φ must map the initial contour S_1 onto the target contour S_2, it can be shown that any displacement field variation V must satisfy

$$V(x)\cdot \mathbf{n}_{S_2}(x + U^*(X)) = 0 \quad \text{for all } x \in S_1 \ . \tag{3.18}$$

Notice that this condition only has to be satisfied precisely on the curve S_1, and that $V = V(x)$ is allowed to vary freely away from the initial contour. The interpretation of the above condition is that the displacement field variation at $x \in S_1$ must be tangent to the target contour S_2 at the point $y = \Phi(x)$. In view of this interpretation of (3.18) it is not difficult to see that necessary condition (3.17) implies that the solution Φ^* of the minimization problem (3.15) must satisfy the following Euler-Lagrange equation:

$$0 = \begin{cases} \Delta U^* - (\Delta U^*\cdot \mathbf{n}^*_{S_2})\,\mathbf{n}^*_{S_2}, & \text{on } S_1, \\ \Delta U^*, & \text{otherwise,} \end{cases} \tag{3.19}$$

where $\mathbf{n}^*_{S_2}(x) = \mathbf{n}_{S_2}(x + U^*(x))$, $x \in S_1$, is the pullback of the normal field of the target contour S_2 to the initial contour S_1. The standard way of solving (3.19) is to use the gradient descent method: Let $U = U(t,x)$ be the time-dependent displacement field which solves the evolution

PDE

$$\frac{\partial U}{\partial t} = \begin{cases} \Delta U - (\Delta U \cdot \mathbf{n}^*_{S_2})\,\mathbf{n}^*_{S_2}, & \text{on } S_1, \\ \Delta U, & \text{otherwise,} \end{cases} \tag{3.20}$$

where the initial displacement $U(0,x) = U_0(x) \in \mathcal{M}$ specified by the user, and $U = 0$ on $\partial\Omega$, the boundary of Ω (Dirichlet boundary condition). Then $U^*(x) = \lim_{t\to\infty} U(t,x)$ is a solution of the Euler-Lagrange equation (3.19).

Notice that the PDE (3.20) coincides with the so-called *geometry-constrained diffusion* introduced in [3]. Thus we have incidentally found a variational formulation of the non-rigid registration problem considered there.

Implementation. Following [3], a time and space discrete algorithm for solving the geometry-constrained diffusion problem can be found by iteratively convolving the displacement field with a Gaussian kernel and then project the deformed contour Γ_1 back onto contour Γ_2 such that the constraints are satisfied (see Algorithm 3). The algorithm needs a initial registration provided by the user. In our implementation we have translated Γ_1 and projected it onto Γ_2 and used this as the initial registration. This gives good results in our case where the deformation and translation is quite small. Dirichlet boundary condition - zero padding in the discrete implementation - have been used. By pre-registration and embedding the image into a larger image, the boundary conditions seems to be a minor practical issue. The displacement field is diffused using convolution in each of x and y coordinates independently with a fix time parameter.

Algorithm 3 The algorithm for the contour matching

INPUT : Contours Γ_1 and Γ_2.
OUTPUT : Displacement field D.

1. **Initial displacement field** Initial registration of the contours.
2. **Diffusion** Convolve the displacement field using a Gaussian kernel.
3. **Deformation** Deform Γ_1 by applying the displacement field D.
4. **Projection** Project the deformed Γ_1 onto Γ_2 (i.e. find the closest point on the contour Γ_2).
5. **Updating the displacement field** Update the displacement field according to matching points on the contour Γ_2
6. **Convergence** Stop if the displacement field is stable, otherwise go to step 2.

3.3.2 Occlusion Detection

The mapping $\Phi = \Phi(x) : \Omega \rightarrow \mathbb{R}^2$ such that Φ maps Γ_1 onto Γ_2 is an estimation of the displacement (motion and deformation) of the boundary of an object between two frames. By finding the displacement of the contour, a consistent displacement of the intensities inside the closed curve Γ_1 can also be found. Φ maps Γ_1 onto Γ_2 and pixels inside Γ_1 are mapped inside Γ_2. This displacement field which only depends on displacement - or registration - of the contour (and not on the image intensities) can then be used to map the intensities inside Γ_1 onto Γ_2. After the mapping, the intensities inside Γ_1 and Γ_2 can be compared and then be classified as the same or different value. Since we can still find the contour in the occluded area, therefore we can also compute the displacement field even in the occluded area.

Implementation. Occlusions are detected by comparing the predicted and the observed intensities inside the segmented object. Unfortunately the displacement field is not exact: it is an estimation of the contour displacement and simultaneously an interpolation of the displacement for pixels inside Γ_1. The intensities in the deformed frame must be interpolated. The interpolation can either be done in the deformed (Lagrange) coordinate or in the original (Euler) coordinate. The next neighbor interpolation scheme in the Euler coordinate has been used. Both the deformed and the current frames are filtered using a low-pass filter to decrease differences due to the interpolation and to the displacement.

The deformed frame, $F_p^{Deformed}(x)$, and the current frame, $F_c(x)$, are compared pixel by pixel using some similarity measures. The absolute differences $|F_p^{Deformed}(x) - F_c(x)|$ are used in our experiments. Different similarity measures require different degree of low-pass filtering. A simple pixel by pixel similarity measure requires more filtering, while a patch based similarity measure may require less or none low-pass filtering. See Algorithm 4.

3.3.3 Re-segmentation

After the occlusion has been detected, the segmentation can be further improved by again employing the previously described Chan-Vese-method augmented with an interaction term. However, in this second stage, the integration is only performed over the area of the image where no occlusion has been detected. This procedure treats the occluded area in the same way as a part of the image with missing data as in [9], which is reasonable. Another way is that the occluded region can be reconstructed in order to improve further the segmentation results. Let Ω_{Occ} be the occlusion masks, e.g. the output after implementing Algorithm 4. Here we

Algorithm 4 The algorithm for occlusion detection using the displacement field to predict the contents in the next frame inside a contour.

INPUT: The previous frame F_p, the current frame F_c, displacement field D
OUTPUT: Occlusion mask.

1. **Deformation** Deform F_p using displacement field D into $F_p^{Deformed}$.
2. **Interpolation** Interpolate $F_p^{Deformed}$ to get in intensities in each grid point.
3. **Low-pass filtering** Low-pass filter the images $F_p^{Deformed}$ and F_c.
4. **Similarity measure** Compare $F_p^{Deformed}$ and F_c inside contour Γ_2 using a similarity measure to get a similarity measure for each pixel.
5. **Thresholding** Find occlusions by thresholding in the similarity measure image.

reconstruct the occluded regions Ω_{Occ} by assigning the intensity values in the occluded regions with the mean value of the intensities inside the contour but excluding the occluded regions, e.g.:

$$I(x) = \mu_{\text{int}}, \forall x \in \Omega_{Occ} \tag{3.21}$$

where

$$\mu_{\text{int}} = \mu_{\text{int}}(\Gamma) = \frac{1}{|\operatorname{int}(\Gamma) \setminus \Omega_{Occ}|} \int_{\operatorname{int}(\Gamma)\setminus\Omega_{Occ}} I(x)\,dx. \tag{3.22}$$

3.4 Experimental Results

Following the Algorithm 2, we implement the proposed model to segment a selected object with approximately uniform intensity frame-by-frame. The minimization of the functional is obtained by the gradient descent procedure (3.12) implemented in the level set framework outlined in Section 3.2.1. Since the Chan-Vese segmentation model finds an optimal piecewise-constant approximation to an image, this model works best in segmenting object that has nearly uniform intensity.

The choice of the coupling constant γ is done manually. It is varied to see the influence of the interaction term on the segmentation results. The contour is only slightly affected by the prior if γ is small. On the other hand, if γ is too large, the contour will be close to a similarity transformed version of the prior. To choose a proper γ is rather problematic in segmentation of image sequences. Using strong prior can give good results when the occlusions occur, but

when segmenting the image frame where occlusions do not occur, the results will be close to the prior.

In Figure 3.2, we show the segmentation results for a nonrigid object in a synthetic image sequence, where occlusion (the gray bar) occurs. Another experiment on a human walking image sequence shown in Figure 3.5 where an occlusion (the superposition of another person) occurs. In both experiments, the standard Chan-Vese method fails to segment the selected object when it reaches the occlusion (Top Row). The result can be improved by adding a frame-to-frame interaction term as proposed in (3.11) (Bottom Row). In these experiments, we use quite large γ to deal with occlusions. As we can see on the last frame in Figure 3.5, the result is close to a similarity transformed of the prior although intensities in between the legs are different from the object.

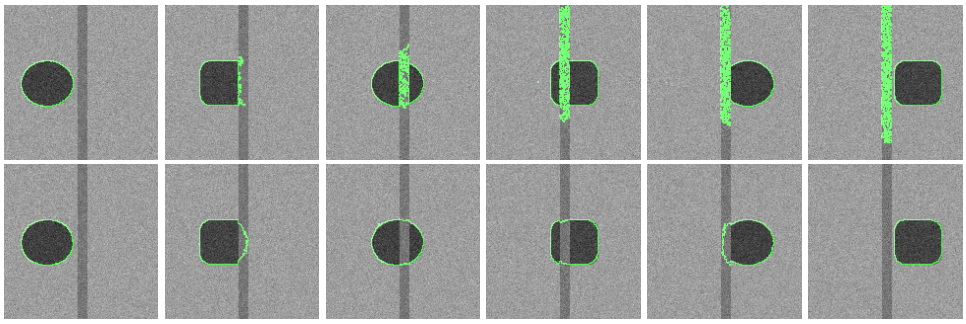


Figure 3.2: Segmentation of a nonrigid object in a synthetic image sequence with additive Gaussian noise (Frame 1-6). Top Row: without the interaction term, noise in the occlusion is captured. Bottom Row: with interaction term, we obtain better results.

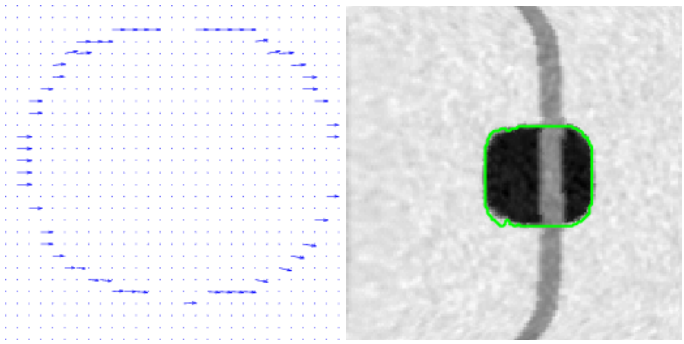


Figure 3.3: Left: Deformation field. Right: Frame 3 after deformation according to the displacement field onto Frame 4.

As described in Section 3.3.1 and Section 3.3, occlusion can be detected and located by deforming the current frame according to the displacement and compare the deformed frame

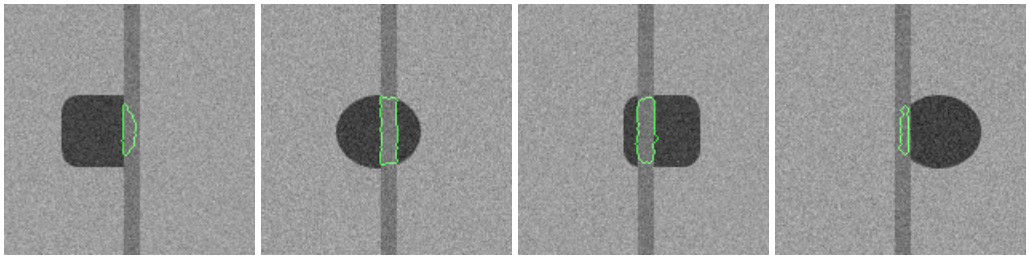


Figure 3.4: Detected occlusions in Frame 2-5 of Figure 3.2.

Figure 3.5: Segmentation of a walking person partly covered by an occlusion in the human walking sequence (Frame 1-5). Top Row: without interaction term, and Bottom Row: with interaction term

Figure 3.6: Detected occlusion in Frame 2-3 of Figure 3.5.

with the next frame (inside the contour S_2). By using the segmentation results of the image sequences, we implement the Algorithm 3 and 4. First we compute the displacement field based on the segmentation results of two frames. In Figure 3.3, we show the displacement field of Frame 3 and 4. With this displacement field, we can do full deformation of the Frame 3 onto Frame 4 (Figure 3.3 right) and then compare the intensities between Frame 4 and deformed Frame 3. By comparing, we can then classify the intensities as having the same or different value by thresholding. In Figure 3.4 and Figure 3.6, we show the occluded regions in the Frame

2-5 of Figure 3.2 and in the Frame 2-3 of Figure 3.5, respectively.

After we reconstruct the occluded regions, we implement the Algorithm 2 again by using smaller coupling constant γ in order to allow more deformation of the contours. As we can see from Figure 3.7 and Figure 3.8, the results are better if we reconstruct the occluded regions than the ones without reconstruction.

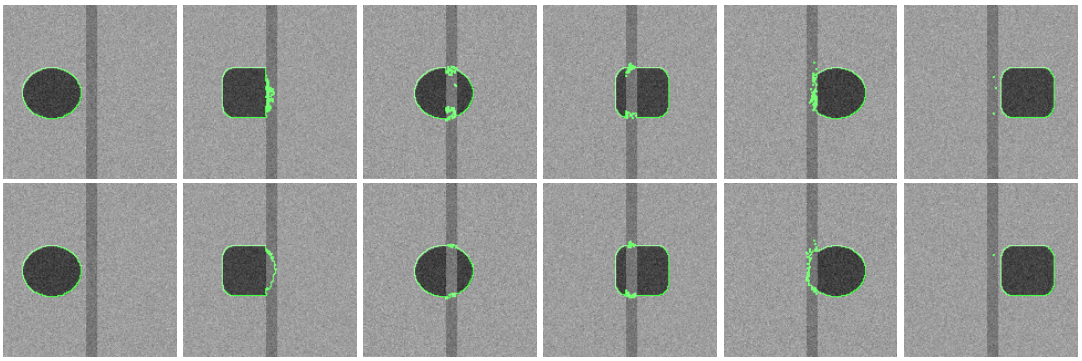


Figure 3.7: Segmentation of the synthetic image sequence by using smaller coupling constant than the one in Figure 3.2. Top row: without reconstruction of the occluded regions. Bottom row: after the occluded regions are reconstructed.

Figure 3.8: Segmentation of the human walking sequence when by using smaller coupling constant than the one in Figure 3.5. Top row: without reconstruction of the occluded regions. Bottom row: after the occluded region is reconstructed.

3.5 Conclusions

We have presented a method for segmentation and occlusion detection of image sequences containing nonrigid, moving objects. The proposed segmentation method is formulated as

variational problem in the level set framework, with one part of the functional corresponding to the Chan-Vese model and another part corresponding to the pose-invariant interaction with a shape prior based on the previous contour. The optimal transformation as well as the shape deformation are determined by minimization of an energy functional using a gradient descent scheme. The segmentation results can then be used to detect the occlusions by the proposed method which is formulated as a variational contour matching problem. By using occlusion information, the segmentation can be further improved by reconstructing the occluded regions. Preliminary results are shown and its performance looks promising.

3. NONRIGID VARIATIONAL OBJECT SEGMENTATION WITH OCCLUSION DETECTION IN IMAGE SEQUENCES

Chapter 4

Continuous Cuts for Prior-Based Object Segmentation

This chapter is a reformatted and slightly modified version of paper [64] without affecting the content or the results. Here we propose a novel prior-based variational object segmentation method in a global minimization framework which unifies image segmentation and image denoising. The idea of the proposed method is to convexify the energy functional of the Chan-Vese method in order to find a global minimizer, so called continuous cuts. The method is extended by adding an additional shape constraint into the convex energy functional in order to segment an object using prior information. We show that the energy functional including a shape prior term can be relaxed from optimization over characteristic functions to optimization over arbitrary functions followed by a thresholding at an arbitrarily chosen level between 0 and 1. Experimental results demonstrate the performance and robustness of the method to segment objects in real images.

4.1 Introduction

Object segmentation is one of the most important and challenging processes in computer vision which aims at extracting objects of interest from a given image. The segmentation results are then used as input for many applications such as recognition, tracking, and classification. The object of interest may exhibit variability in pose and shape which makes segmentation difficult and still a major topic of research.

Many approaches have been proposed to solve the object segmentation problem. In particular, variational methods for image segmentation have had great success such as snakes [83], geodesic active contours [27], geodesic active region [116] and the Chan-Vese method [31]. Yet, the main drawback of those methods is the existence of local minima due to non-convexity of the energy functionals. Minimizing those functionals by gradient descent methods makes the initialization critical. A number of methods have been proposed to find global minima such as [4, 20, 33]. Their approaches give promising results, but it is unclear how to integrate shape constraints.

Integrating shape priors into active contour methods has been proven to improve the robustness of the segmentation methods in the presence of occlusions, clutter, and noise in the images. Various methods have been proposed to address shape prior integration into segmentation process such as [35, 41, 93, 126, 129, 146] and the references therein.

This chapter suggests a novel variational approach to prior-based segmentation by adding a shape prior into the global minimization framework using the Mumford-Shah [106] and the Chan-Vese [31] models. The segmentation process is carried out concurrently with the denoising of the image and the transformation of the shape prior to the object of interest. The idea of the proposed method is to use the relaxation of the non-convex energy functional of the Chan-Vese model to the minimization over all functions in such a way that the minimizer of the extended functional can be transformed into the minimizer for the original model by simple thresholding as done in [33]. This method is often called continuous cuts and the relaxed energy functional can then be minimized by gradient descent methods to find a global minimum. The main contribution of this chapter is to extend this method to also include a shape prior term while maintaining the relaxation property.

4.2 Continuous Cuts

Minimizing the variational formulation of the Chan-Vese method [31] by gradient descent methods can get stuck in local minima due to the non-convexity of the energy functional. In order to overcome this, Chan *et al.* [33] propose to convexify the energy functional of the Chan-Vese method [31]. By introducing an auxiliary variable u, the Chan-Vese method can be

reformulated as the following non-convex minimization problem

$$\min_{u=\mathbf{1}_{\Sigma(x)}} \Big\{ E_s(u) = \int_\Omega |\nabla u| \\ +\lambda \int_\Omega \Big\{ \big(I(x)-\mu_1\big)^2 - \big(I(x)-\mu_2\big)^2 \Big\} u\,dx \Big\}, \tag{4.1}$$

where $\mathbf{1}_{\Sigma(x)}$ denotes a characteristic function of a subset Σ of Ω and $\lambda, \mu_1, \mu_2 \in \mathbb{R}$ and $I(x)$ is the given image. In the next step (4.1) is relaxed to the convex problem

$$\min_{0\le u\le 1} E_s(u), \tag{4.2}$$

where now u is an arbitrary function bounded between 0 and 1. If $u(x)$ is a minimizer of (4.2), then the set $\Gamma(\mu) = \{x \in \Omega, u(x) > \mu\}$ has to be a minimizer of the Mumford-Shah functional [106] for a.e. $\mu \in [0,1]$, implying that the solution to (4.1) can be obtained by thresholding u at an arbitrary threshold between 0 and 1, for details see [33]. By having a convex energy functional, we can get a global minimum by using a standard gradient descent method. Notice that (4.2) is not strictly convex which means that it can have several global minima.

4.3 Shape Priors for Continuous Cuts

We would now like to introduce an additional shape prior term into (4.1) and the natural choice is to use a shape prior energy of the form

$$E_p(u) = \int_\Omega (u - \mathbf{1}_{\Omega_p})^2\,dx, \tag{4.3}$$

where $\mathbf{1}_{\Omega_p}$ is the characteristic function of the shape prior template. Inspired by the fact that $\nabla E_p(u) = 2(u - \mathbf{1}_{\Omega_p})$, we consider the minimization problem

$$\min_{u=\mathbf{1}_{\Sigma(x)}} \Big\{ E_{sp}(u) \\ = E_s(u) + \gamma \int_\Omega \big(\hat{u} - \mathbf{1}_{\Omega_p}(x)\big)\, u\,dx \Big\}, \tag{4.4}$$

where $\gamma \in \mathbb{R}$ and $\hat{u}$ is a 'frozen' u which is updated after finding a solution to (4.4). We further relax (4.4) to

$$\min_{0\le u\le 1} E_{sp}(u). \tag{4.5}$$

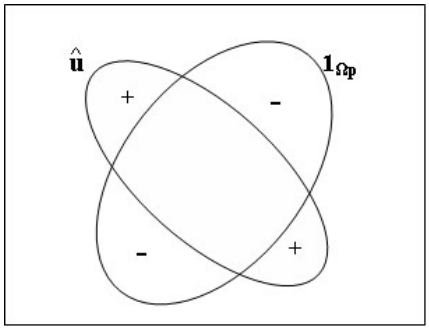


Figure 4.1: Motivation for using the shape prior as in (4.4)

Note that (4.5) still preserves the convexity of (4.2) with respect to u. The reason for not using u directly is that we would like to preserve the property that the solution to (4.4) can be obtained from the solution to (4.5) by a simple thresholding. Also note that the shape prior term $\left(\hat{u}-\mathbf{1}_{\Omega_p}(x)\right)$ has the property that it is positive on $\{\hat{u}\geq\mu\}\backslash\Omega_p$ pushing u to zero and negative on $\Omega_p\backslash\{\hat{u}\geq\mu\}$ pushing u to one, which is driving the set $\{\hat{u}\geq\mu\}$ towards Ω_p, see Figure 4.1. We are now ready to prove the main theorem of this chapter:

Theorem 4.3.1 *For any given $\mu_1,\mu_2\in\mathbb{R}$ and $\hat{u}\in\mathbb{R}^2$, a global minimizer of* (4.4) *is also a global minimizer of* (4.5).

Proof We use the coarea formula and the proof in [33] with additional shape prior term

$$\begin{aligned}&\int_\Omega(\hat{u}-\mathbf{1}_{\Omega_p})\,u(x)\,dx\\ =&\int_0^1\int_{\Omega\cap\{x:u(x)>\mu\}}(\hat{u}-\mathbf{1}_{\Omega_p})\,dx\,d\mu.\end{aligned}$$

Setting $\Gamma(\mu):=\{x:u(x)>\mu\}$, for any $u(x)\in L^2(\Omega)$ such that $0\leq u(x)\leq 1$ for all $x\in\Omega$, we have (4.5) is equal to

$$\begin{aligned}&\min_{u,\mu_1,\mu_2}\left\{\int_0^1\left\{\mathrm{Per}(\Gamma(\mu);\Omega)\right.\right.\\ &+\lambda\int_{\Gamma(\mu)}(I(x)-\mu_1)^2\,dx+\lambda\int_{\Omega\backslash\Gamma(\mu)}(I(x)-\mu_2)^2\,dx\\ &\left.\left.+\gamma\int_{\Gamma(\mu)}(\hat{u}-\mathbf{1}_{\Omega_p})\,dx\right\}d\mu-C\right\},\end{aligned}$$

where $\int_0^1\mathrm{Per}(\Gamma(\mu);\Omega)\,d\mu=\int_\Omega|\nabla u|$ and $C=\int_\Omega(I(x)-\mu_2)^2\,dx$ is independent of u. It follows

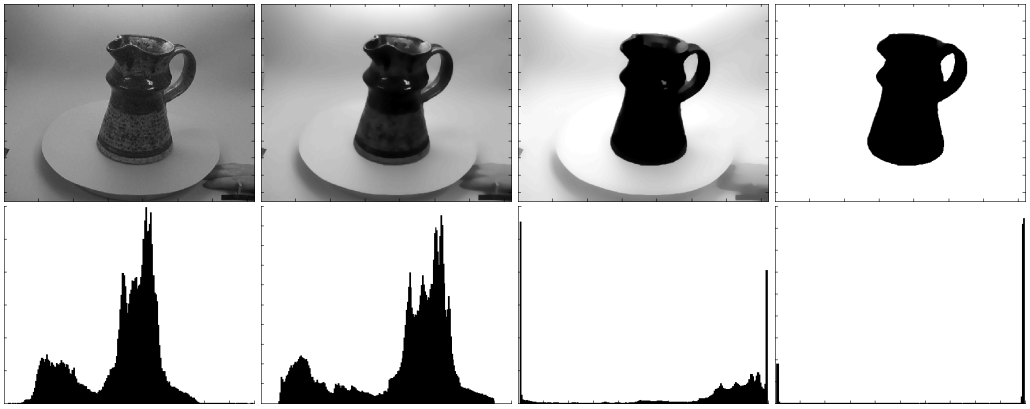


Figure 4.2: The evolution of $u(x)$ (top row) and the corresponding histogram (bottom row). First column: initial. Middle columns: intermediate results. Right column: final results

that if $u(x)$ is a minimizer of (4.5), then it is also a minimizer of (4.4). ∎

Corollary 4.3.2 *The solution to (4.4) can be obtained from the solution to (4.5) by thresholding at an arbitrary level between* 0 *and* 1.

In order to make (4.5) invariant with respect to similarity transformations, the convex minimization problem of (4.5) is reformulated by adding transformation parameters, as in [35], with the respect to the shape prior $\mathbf{1}_{\Omega_p}$

$$\min_{0\leq u\leq 1,s,\theta,\mathbf{a}} \left\{ E_s + \gamma \int_\Omega \left(\hat{u} - \mathbf{1}_{\Omega_p}(sRx+\mathbf{a}) \right) u \right\}, \tag{4.6}$$

for the scaling s, translation vector $\mathbf{a}$, and rotation matrix $R(\theta)$:

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Notice that the minimization problem of (4.6) is no longer convex in all unknowns, but the convexity with respect to u facilitates the computation of the transformation parameters. To minimize (4.5), the constrained minimization problem is reformulated as the unconstrained minimization problem by adding a penalty term $v(u)$

$$\min_u \left\{ E^e_{sp}(u) = E_{sp}(u) + \alpha \int_\Omega v(u) \right\}, \tag{4.7}$$

where $v(\xi) := \max\{0, 2|\xi - \frac{1}{2}| - 1\}$ and $\alpha > \frac{\lambda}{2} \|(I(x)-\mu_1)^2 - (I(x)-\mu_2)^2\|_{L^\infty(\Omega)}$. This procedure is exactly the one used in [33] and it can be proven in the same way that a solution to (4.7)

is also a solution to (4.5). The function $v(\xi)$ is then regularized as $v_\varepsilon(\xi)$ by a small constant $\varepsilon > 0$ to smooth the sharp bend at 0 and 1. We choose regularized v as the following:

$$v_\varepsilon(\xi) = \begin{cases} -\xi & \text{if} \quad \xi < -\varepsilon, \\ -\xi + 25(\varepsilon+\xi)^2 & \text{if} \quad -\varepsilon \le \xi < \varepsilon, \\ 0 & \text{if} \quad \varepsilon \le \xi < 1-\varepsilon, \\ -1+\xi+25(\varepsilon+1-\xi)^2 & \text{if} \quad 1-\varepsilon \le \xi < 1+\varepsilon, \\ \xi - 1 & \text{if} \quad 1+\varepsilon \le \xi, \end{cases}$$

where $\varepsilon > 0$ is a small constant. In [33], it is proven that by adding a penalty term v (4.6) has the same set of minimizers as (4.7) (without the shape prior term $\gamma \int_\Omega \left(\hat{u} - \mathbf{1}_{\Omega_p}(sRx+\mathbf{a})\right) u\,dx$).

4.4 Implementation and Results

The proposed energy functional (4.6) is minimized with respect to u and the transformation parameters s, R, a by gradient descent method. The evolution equations associated with the Euler-Lagrange equations of (4.6) are the following

$$\begin{aligned} \frac{\partial u}{\partial t} &= \nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right) - \alpha v_\varepsilon'(u) \\ &\quad -\lambda\left(\left(I(x)-\mu_1\right)^2 - \left(I(x)-\mu_2\right)^2\right) \\ &\quad -\gamma\left(\hat{u} - \mathbf{1}_{\Omega_p}(sRx+\mathbf{a})\right), \end{aligned} \tag{4.8}$$

$$\frac{\partial s}{\partial t} = \gamma \int_\Omega \nabla \mathbf{1}_{\Omega_p}(sRx+\mathbf{a})\, uRx\,dx, \tag{4.9}$$

$$\frac{\partial \theta}{\partial t} = \gamma \int_\Omega \nabla \mathbf{1}_{\Omega_p}(sRx+\mathbf{a})\, us\frac{dR}{d\theta}x\,dx, \tag{4.10}$$

$$\frac{\partial \mathbf{a}}{\partial t} = \gamma \int_\Omega \nabla \mathbf{1}_{\Omega_p}(sRx+\mathbf{a})\, u\,dx, \tag{4.11}$$

where $v_\varepsilon'(u)$ denotes the derivative of $v_\varepsilon(u)$. Here the gradient descent of u (4.8) is coupled with gradient descents of the transformation parameters which update dynamically the transformation parameters to map $\mathbf{1}_{\Omega_p}$ and $\hat{u}$ in the best possible way (see Algorithm 5).

We implement the proposed method to segment objects in images. As shown in Figure 4.2, $u(x)$ takes values between 0 and 1 during the evolution and at convergence, it is very close to being binary. The values of $u(x)$ at the end accumulate near 0 and 1, as shown in the histograms of $u(x)$. The regularized exact penalty term $v_\varepsilon(\xi)$ in (4.7) prevents them to be 0 and 1. Figure 4.3 and 4.4 show the segmentation results of a bird [97] and a cup. The given images are used

Algorithm 5 Algorithm for minimizing the proposed segmentation functional

INPUT: $I, u, \mathbf{1}_{\Omega_p}, s, \theta, \mathbf{a}, \alpha, \lambda, \gamma, \Delta t$
OUTPUT: Optimal u

1. Compute μ_1 and μ_2 as mean intensities of region inside and outside the contour.
2. Compute the transformation parameters using (4.9), (4.10), and (4.11).
3. Transform the prior.
4. Update u using (4.8).
5. Repeat until convergence.

as the initial of $u(x)$ and the contours are represented by $u(x) = 0.5$. As we can see, at the convergence state, global minima are found and the segmentation results can then be improved to segment objects of interest by adding shape priors, which are segmented manually and are then transformed to different pose, despite the presence of occlusions.

4.5 Conclusions

We have proposed a novel variational region-based active contour method for prior-based object segmentation in a global minimization framework. The method is based on convexifying the energy functional of Chan-Vese method and adding a shape prior term as a constraint to segment an object whose global shape is given. The energy functional can be relaxed from optimization over characteristic functions to over arbitrary functions followed by a thresholding at an arbitrarily chosen level between 0 and 1. Experimental results demonstrate desirable performance of the method to segment objects of interest in the images.

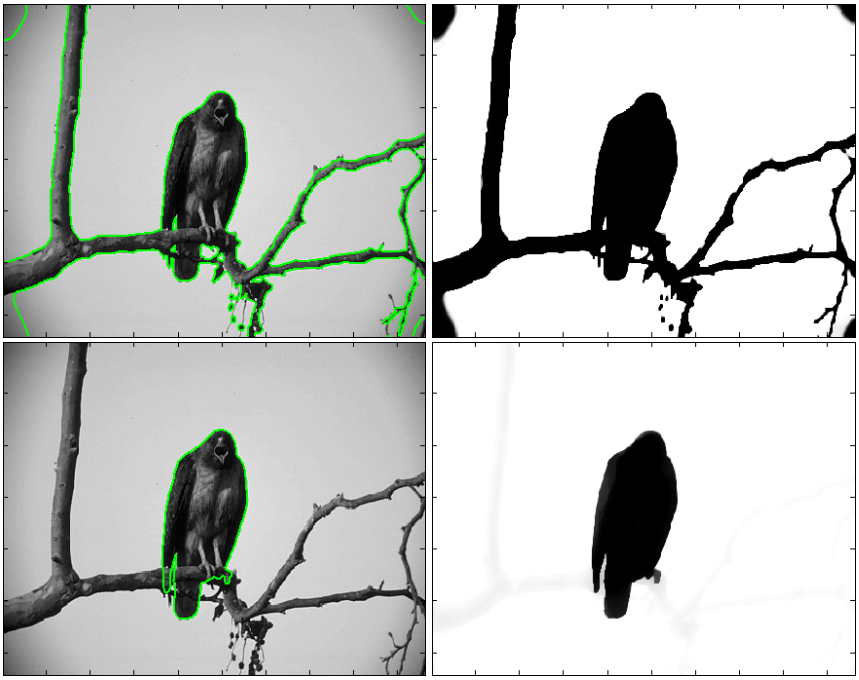


Figure 4.3: Segmentation of a bird using continuous graph cuts and its corresponding u. Top row: without a shape prior. Bottom row: with a shape prior.

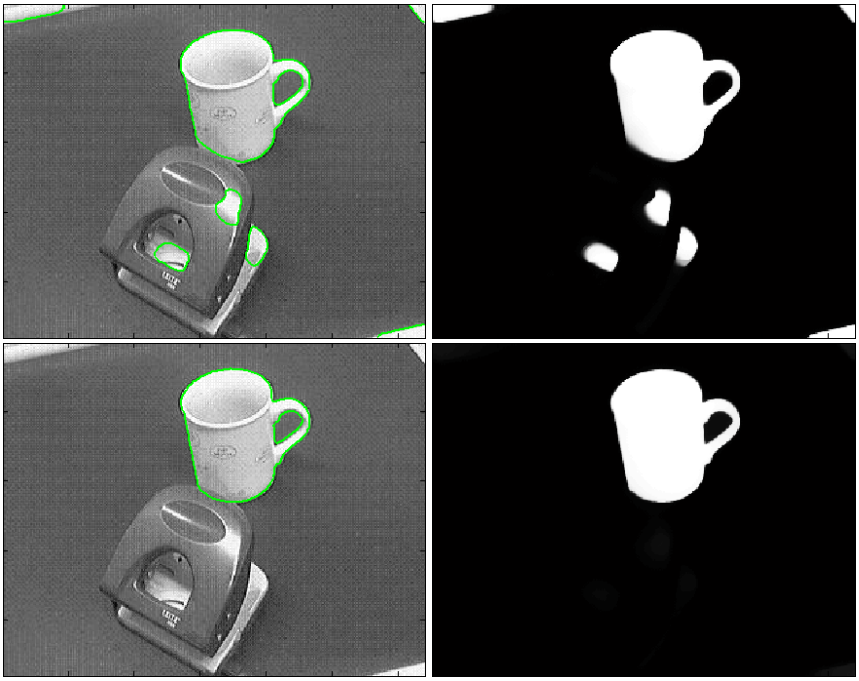


Figure 4.4: Segmentation of a cup using continuous graph cuts and its corresponding u. Top row: without a shape prior. Bottom row: with a shape prior.

Chapter 5

View Point Tracking of Rigid Objects using Region-Based Segmentation Model and Shape Priors from Shape Submanifolds

This chapter is a reformatted and slightly modified version of paper [73] without affecting the content or the results. Here we study the task to infer and to track the viewpoint onto a 3D rigid object by observing its image contours in a sequence of images. To this end, we consider the manifold of invariant planar contours and learn the low-dimensional submanifold corresponding to the object contours by observing the object off-line from a number of different view-points. This submanifold of object contours can be parametrized by the view sphere and, in turn, be used for keeping track of the object orientation relative to the observer, through interpolating samples on the submanifold in a geometrically proper way. Our approach replaces explicit 3D object models by the corresponding invariant shape submanifolds that are learnt from a sufficiently large number of image contours, and is applicable to arbitrary objects.

5.1 Introduction

The representation of planar shapes has been a focus of research during the last few years [85, 101, 102, 139]. By mathematically separating similarity transforms and potentially also

reparameterizations from other deformations of planar curves, an invariant representation of planar *shapes* is obtained in terms of a smooth manifold embedded in a Euclidean space. Furthermore, distances between shapes can be computed that are only sensitive to shape deformations, by determining the geodesic path between the corresponding points of the shape manifold (Fig. 5.3 below provides an illustration).

Here we adopt this representation and show that it is accurate enough to infer the change in aspect of a given rigid 3D object, represented by a point on the view sphere, just by observing 2D shapes of its silhouette in a given image sequence – see the left panel of Fig. 5.1 below.

To this end, we assume to be given a collection of silhouettes of any known object, that we represent one-to-one by a corresponding set of points on the view sphere. These data can be acquired off-line by observing the object from different directions. We regard these shapes as samples of an object-specific *submanifold* of the manifold of all planar shapes, that is parameterized by the view sphere. Taking into account the geometry of this submanifold and interpolating the shape samples accordingly, we show that either the viewpoint of a moving camera, or object pose relative to the observer, can be tracked by observing deformations of the object's silhouette in an image sequence.

We point out the *3D models are not utilized* in this work, besides illustrating graphically various points below. Rather, a sample set of object contours observed from different viewpoints, along with the information to what object they belong, define the input data. Our results are novel and relevant, for instance, for reaching and maintaining a reference position relative to a moving object, through vision based control, in man-made and industrial scenes.

Related work. Related work has been published recently in [47, 58, 82, 90]. Etyngier et al. [58] use Laplacian eigenmaps [10] for embedding a set of training shapes into a low dimensional standard Euclidean space. They present a method for projecting novel shapes to the submanifold representing the training samples, in order to model a shape prior for image segmentation. Similarly, Lee and Elgammal [90] use locally linear embedding (LLE) [131] to learn separately a configuration manifold of human gaits and a view manifold corresponding to a circle on the view sphere, based on a tensor factorization of the input data.

While nonlinear Euclidean embeddings (Laplacian eigenmap, LLE) of locally connected similarity structures (weighted adjacency graphs) are employed in [58, 90], we use directly the intrinsic manifold of invariant shapes as developed in [85, 101]. Statistical models based on this manifold have been elaborated in [47, 82] for *deformable* objects and shapes of *classes* of rigid objects, respectively, in connection with image segmentation and computational anatomy.

By contrast, we focus on *tracking and pose estimation* of a single rigid object, based on contour deformations and the corresponding shape submanifold. This approach is novel. Our work may be regarded as a *learning-based* approach for associating views and contours of *arbitrary* objects, that is both more general and easier to apply than earlier work on *model-based* contour representations of *specific* objects [57, 148].

5.2 Shape Model, Object Representation, Learning

We work with the *elastic closed pre-shape space* covering closed regular two-dimensional curves, proposed in [101]. A regular curve $\alpha : [0,1] \mapsto \mathbb{R}^2$ is represented by $\alpha(t) = \alpha_0 + \int_0^t e^{\Phi(t)} e^{i\Theta(t)}\, dt$, with the integrand denoting a velocity function along the curve. $e^{\Phi(t)}$ describes the curve speed, while $\Theta(t)$ is the tangent angle relative to the real axis in the complex plane. To achieve invariance under translation, the integral constant α_0 is left out, and shapes are represented by pairs (Φ,Θ) as elements of a vector space of functions denoted by $\mathcal{H}$. To also achieve scale and rotation invariance and to restrict to closed regular curves, further constraints turn $\mathcal{H}$ into the space of pre-shapes $\mathcal{C}$:

$$\mathcal{C} := \left\{ (\Phi,\Theta) \in \mathcal{H} : \begin{array}{ll} \int_0^1 e^{\Phi(t)} e^{i\Theta(t)}\, dt = 0 & \text{(closure)} \\ \int_0^1 e^{\Phi(t)}\, dt = 1 & \text{(scale)} \\ \int_0^1 \Theta(t) e^{\Phi(t)}\, dt = \pi & \text{(rotation)} \end{array} \right\}. \tag{5.1}$$

So, curves are restricted to length 1 and an angle function average of π. Note that this is an arbitrary choice, we adopted π from [101]. Invariance with respect to reparameterizations is not handled intrinsically, since it would raise a considerable additional computational burden. Instead, shapes are matched by dynamic programming, following [133].
The elastic Riemannian metric [151] used on $\mathcal{C}$ is

$$\langle (p_1,t_1),(p_2,t_2) \rangle_{(\Phi,\Theta)} := a \int_0^1 p_1(t) p_2(t) e^{\Phi(t)}\, dt + b \int_0^1 t_1(t) t_2(t) e^{\Phi(t)}\, dt \tag{5.2}$$

with constants $a,b \in \mathbb{R}$ that weight the amount of stretching and bending, and with tangent vectors $(p_{\{1,2\}}, t_{\{1,2\}})$ at (Φ,Θ). [101] proposes ways to numerically approximate geodesics on a discrete representation of $\mathcal{C}$, as well as to approximate the inverse exponential map by gradient descent on $\mathcal{C}$. Another recent representation of elastic shape is discussed in [80], also cf. [100], which allows for faster computations. However, rotation invariance is not easy to

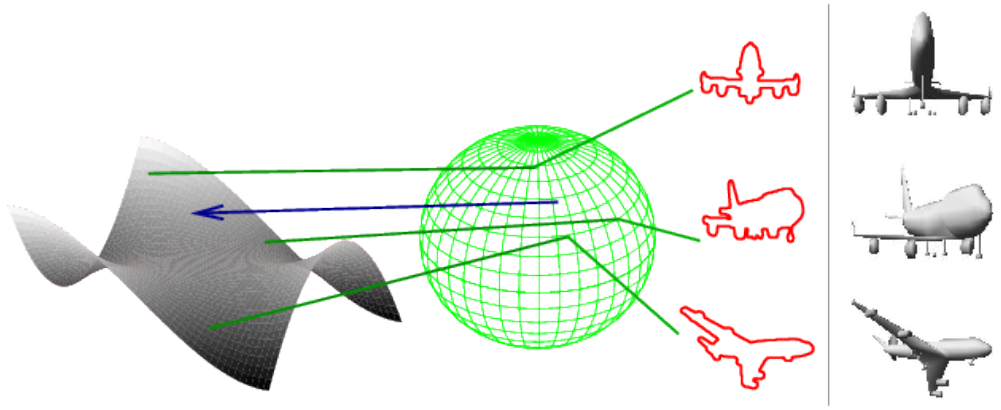


Figure 5.1: Illustration of a view sphere. Right hand: indicated are three sampled contours of an airplane seen from a camera from points on the view sphere. The object is located in the center of the sphere. Left hand: illustration of the shape sub-manifold. The green lines between sphere and manifold indicate corresponding points, the blue arrow indicates a point that is interpolated using, in this case, three points which are neighbors on the sphere. This specific object was taken from the Princeton 3D shape benchmark [135].

achieve. [81] introduces an optimization approach to find minimal geodesics between orbits under the action of rotations and reparameterizations.

View Sphere Sampling. The input data of our approach are given samples on the view sphere $\mathbb{S}^2$ from any object, at known positions (see Fig. 5.1). These data are acquired off-line and result in a sample set of points in $\mathcal{C}$.

5.3 Pose Inference and Tracking on the View Sphere

This section describes a model that we use for modeling motion of a point on the sphere that represents the object's shape in a submanifold of $\mathcal{C}$, as well as a simple scheme to predict positions locally. We also explain how we keep track of points on the view sphere that correspond to shapes measured from images in an image sequence. To avoid confusion with tracking an object *in the image plane*, we call the process of tracking the position on the view sphere *sphere tracking*.

Motion Model. We model a mass point on the sphere as motion in a potential field $V(x) = \mathfrak{m} \cdot g \cdot (x-P)^2$, together with a friction component. $\mathfrak{m}$ is a constant inertia, g weights the impact of V, and β in Equation (5.3) weights the impact of friction. The motion is governed by the

differential equation

$$\underbrace{-2 \cdot \mathfrak{m} \cdot g \cdot (s(t) - P)}_{-\nabla V} \underbrace{-\beta \cdot \dot{s}(t)}_{\text{Stokes friction}} = \mathfrak{m} \cdot \ddot{s}(t)\,. \tag{5.3}$$

This is applied to a point in the tangent space of the group of 3D rotations, i.e. $s(t), P \in T\mathcal{SO}_3$, with rotations representing motions of a point on the sphere $\mathbb{S}^2$. The corresponding exponential and logarithmic maps for $\mathcal{SO}_3$ can be efficiently computed in closed form. The "center of gravitation" P is updated whenever a new measurement P_k is available. Fig. 5.2 shows an illustration of the motion model following a path of points P.

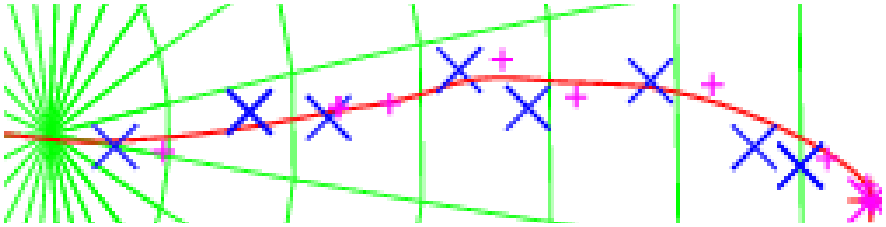


Figure 5.2: Representing and tracking shape changes as motions on the view sphere. Blue: measurements P_k. Red: path $s(t)$ of the mass point. Magenta: predicted points. The start point of the trajectory is at the far left end. The green grid lines indicate the underlying sphere.

Predictions. Given past measurements $p_i \in \mathbb{S}^2$, we would like to predict $s(t)$ locally. Assume to be given a new measurement P_k at time t_k, and the motion model to be at point $s(t_k)$. We then follow the trajectory governed by (5.3) until the distance $d(s(t_k), P_k)$ has been traveled, say at time t'_k, so that $d(s(t_k), s(t'_k)) = d(s(t_k), P_k)$, and then continue for an additional fixed time period $\Delta t = t'_k - t_k$ to obtain the prediction

$$p_{pred} := s(t'_k + \Delta t)\,. \tag{5.4}$$

As illustrated in Fig. 5.2, this simple "mechanical" model can result in rather sensible paths and corresponding predictions of shape changes, as detailed below.

Shape Interpolation. Interpolation of shapes on the view sphere at a point $s \in \mathbb{S}^2$ is realized by Karcher means using a local neighborhood M of sampled shapes around s. The empirical Karcher mean is

$$\mu = \arg\min_{m \in \mathcal{C}} \sum_{i=1}^{|M|} a_i \cdot d(m, p_i)^2\,, \tag{5.5}$$

with $d(\cdot,\cdot)$ the geodesic distance on $\mathcal{C}$, and weights $a_i \geq 0$ with $\sum_i a_i = 1$. μ can in practice

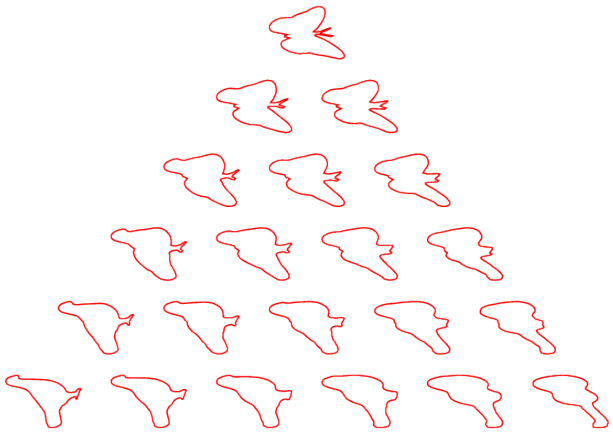


Figure 5.3: Illustration of shape interpolation with Karcher means in the closed pre-shape space $\mathcal{C}$. The corners represent the original shapes, the other contours are interpolations weighted with their barycentric coordinates. The corner curves are randomly chosen from the MPEG-7-CE1 shape data base.

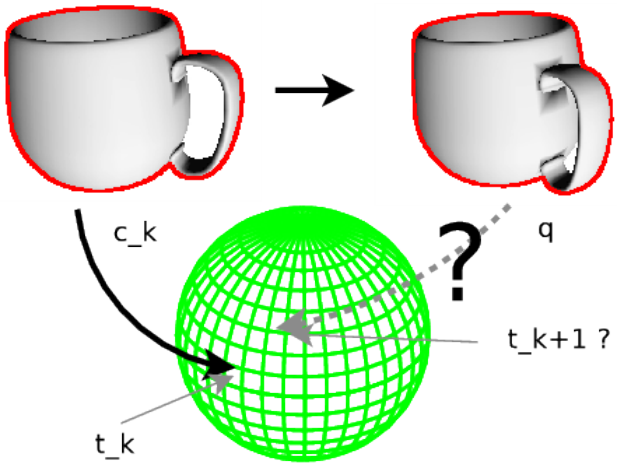


Figure 5.4: Keeping track of the spherical position: Shape c_k and position t_k are known, as well as a new shape q. What is the (approximate) position t_{k+1} on the view sphere corresponding to q?

be calculated by gradient descent [118]. Fig. 5.3 illustrates the interpolation of three shapes depicted at the corners of the triangle.

Keeping Track of the Spherical Position. Assume that we know initially a point $c_k \in \mathcal{C}$ and the corresponding position $t_k \in \mathbb{S}^2$. Now, suppose a new shape $q \in \mathcal{C}$ is to be considered, typically delivered by an image segmentation algorithm that tracks an object over a number of frames (see the next section). Fig. 5.4 illustrates the following problem: We wish to determine a point $c_{k+1} \in \mathcal{C}$ at $t_{k+1} \in \mathbb{S}^2$ on the sub-manifold modeled by the samples p_i from the view sphere at spherical coordinates $t_i \in \mathbb{S}^2$, that is as close as possible to q. That is, we would like to minimize the geodesic distance $d(m,q) = \|\mathrm{Log}_m(q)\|_m$ by minimizing

$$F(m,q) = \|\mathrm{Log}_m(q)\|_m^2, \tag{5.6}$$

where m results from minimizing (5.5),

$$m(t) = \arg\min_{\tilde{m} \in \mathcal{C}} \left(\sum_{i=1}^{|M|} a_i(t) \cdot d(\tilde{m}, p_i)^2 \right) \tag{5.7}$$

with both the neighborhood M and the weights a_i depending on the spherical position t. We then solve at frame $k+1$

$$t^\star_{k+1} = \arg\min_t F(m(t), q) \tag{5.8}$$

using non-linear conjugate gradient descent on the view sphere, as follows: choose $b_{\ell,1}, b_{\ell,2} \in \mathbb{R}^3$ to be orthonormal basis vectors of the tangent space $T_{t^\ell}(\mathbb{S}^2)$, and a small constant $\Delta > 0$. Notice that in the following equations, Exp and Log denote the exponential and inverse exponential maps on the sphere $\mathbb{S}^2$, not on the pre-shape space $\mathcal{C}$.

$$\text{trans} : T(\mathbb{S}^2) \times \mathbb{S}^2 \times \mathbb{S}^2 \mapsto T(\mathbb{S}^2), \quad v_2 = \text{trans}(v_1, t_1, t_2) \tag{5.9}$$

is a function that takes a tangent vector at t_1 and translates it along a geodesic from t_1 to t_2. Then, let $t^0 = t^\star_k$, $\beta_{-1} = 0$, $\tilde{d}_{-1} = 0$, and

$$v_\ell = \sum_{i=1}^{2} b_{\ell,i} \cdot \frac{F(m(\text{Exp}_{t^\ell}(\Delta \cdot b_{\ell,i})), q) - F(m(t^\ell), q)}{\Delta} \tag{5.10}$$

$$d_\ell = -v_\ell + \beta_{\ell-1} \tilde{d}_{\ell-1} \tag{5.11}$$

$$t^{\ell+1} = \text{Exp}_{t^\ell}(\alpha \cdot d_\ell) \tag{5.12}$$

$$\tilde{d}_\ell = \text{trans}(d_\ell, t^\ell, t^{\ell+1}) \tag{5.13}$$

$$\tilde{v}_\ell = \text{trans}(v_\ell, t^\ell, t^{\ell+1}) \tag{5.14}$$

$$\beta_\ell = \frac{[v_{\ell+1} - \tilde{v}_\ell]^\top v_{\ell+1}}{v_\ell^\top v_\ell}. \tag{5.15}$$

v_ℓ takes the role of the gradient direction, in the tangent space of $\mathbb{S}^2$ at the current point t^ℓ. d_ℓ is the search direction, computed from the gradient v_ℓ and the previous search direction $\tilde{d}_{\ell-1}$, with factor $\beta_{\ell-1}$ calculated using the Polak-Ribière variant of the non-linear conjugate gradient method in Equation (5.15), which is more robust than the Fletcher-Reeves variant according to [109]. The rest of the above equations are needed to adapt to the geometry of the sphere. Specifically, we need to translate tangent vectors to the current iterate t^ℓ to be able to combine them, and we need to go back to the sphere using the exponential map.

In order to find a step length $\mathbb{R} \ni \alpha > 0$ for use in Equation (5.12), we use a standard line

search procedure with the Armijo or *sufficient decrease* condition

$$F(m(\mathrm{Exp}_{t^\ell}(\alpha \cdot d_\ell)), q) \leq F(m(t^\ell), q) + c \cdot \alpha \cdot (v_\ell^\top d_\ell), \quad 0 < c < 1\,. \tag{5.16}$$

Figures 5.5 and 5.6 depict paths on the view sphere.

5.4 Segmentation and Image Contours

There are several possibilities to obtain contours from actual images, and to track contours while they are moving *in the image plane*. We have so far applied two methods: the well known region-based segmentation method of the Chan-Vese [31] and the related, more recent method from [33]. Since [33] finds a global optimum and is suitable if the images contain only a more or less homogeneous background and a single object. In more complex scenes containing clutter and heterogeneous background the level set method that only finds local optima is advantageous. We sketch these two approaches below, and how results from the sphere tracking are used as prior for steering the segmentation process.

Level sets. Our implementation of level set segmentation uses the image energy from [31] and additionally the curvature diffusion regularization term from [51], replacing the more common mean curvature term in the evolution in all our experiments. We also optionally use a prior energy based on [35] and [127]:

$$E_{Shape} = \int_\Omega \left[H(\phi(x)) - H(\phi_0(sRx + a))\right]^2 dx. \tag{5.17}$$

H denotes the Heaviside function, ϕ and ϕ_0 are the embedding functions of the evolving contour and the prior contour, respectively and the scale $s \in \mathbb{R}$, translation $\mathbf{a} \in \mathbb{R}^2$, and rotation matrix

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Differentiating (5.17) with respect to the transformation parameters yields

$$\begin{aligned}
\frac{dE}{ds} &= -\int_\Omega Rx\left[H(\phi(x)) - H(\phi_0(sRx+\mathbf{a}))\right] \cdot H'(\phi_0(sRx+\mathbf{a}) \cdot \nabla\phi_0(sRx+\mathbf{a})\, dx, \\
\frac{dE}{d\theta} &= -\int_\Omega \left[H(\phi(x)) - H(\phi_0(sRx+\mathbf{a}))\right] \cdot H'(\phi_0(sRx+\mathbf{a}) \cdot s \cdot \frac{dR}{d\theta} \cdot \nabla\phi_0(sRx+\mathbf{a})\, dx, \\
\frac{dE}{d\mathbf{a}} &= -\int_\Omega \left[H(\phi(x)) - H(\phi_0(sRx+\mathbf{a}))\right] \cdot H'(\phi_0(sRx+\mathbf{a}) \cdot \nabla\phi_0(sRx+a)\, dx,
\end{aligned}$$

which we use in a gradient descent to solve $\min_{s,\theta,\mathbf{a}} E_{Shape}$.

Global Segmentation Method. The variational segmentation model of [31] suffers from the existence of local minima due to the non-convexity of the energy functional. Segmentation results depend on the initialization. To overcome this limitation, Chan *et al.* [33] propose algorithms which are guaranteed to find global optima as follows: For a normalized grey scale image $I(x) : \Omega \mapsto [0,1]$ on the domain Ω and constants $\lambda, \mu_1, \mu_2 \in \mathbb{R}$, a global minimizer u can be found by minimizing the *convex* functional

$$\min_{0\leq u\leq 1} \int_\Omega |\nabla u|\, dx + \lambda \int_\Omega \{(I(x)-\mu_1)^2 - (I(x)-\mu_2)^2\}\, u(x)\, dx\,. \tag{5.18}$$

As mentioned and proven in [33] that if $u(x)$ is a solution of (5.18), then for almost every $\gamma \in [0,1]$ $\mathbf{1}\{x : u(x) > \gamma\}(x)$ is a global minimizer of [31].

In order to segment an object of interest in the image plane, we modify (5.18) by adding an additional term as shape prior

$$\min_{0\leq u\leq 1} \int_\Omega |\nabla u|\, dx + \lambda \int_D \{(I(x)-\mu_1)^2 - (I(x)-\mu_2)^2 + (\hat{u}(x)-\tilde{u}(x))\}\, u(x)\, dx\,, \tag{5.19}$$

where $\hat{u}$ is a 'frozen' u which gets updated after each time step in the numerical implementation, and $\tilde{u}$ is the prior template. We would like (5.19) to be invariant with respect to Euclidean transformations of the object in the 2D image plane. To this end, we add transformation parameters, as in [35], of the fixed $\hat{u}$ with respect to the prior $\tilde{u}$ by minimizing

$$E_{Shape} = \int_\Omega [\hat{u}(x) - \tilde{u}(sRx+\mathbf{a})] u(x)\, dx. \tag{5.20}$$

As a result, we obtain

$$\min_{u,s,R,\mathbf{a}} \int_\Omega |\nabla u|\, dx + \lambda \int_\Omega \{(I(x)-\mu_1)^2 - (I(x)-\mu_2)^2 + (\hat{u}(x)-\tilde{u}(sRx+\mathbf{a}))\}\, u(x)\, dx\,, \tag{5.21}$$

which is minimized by gradient descent. This functional is no longer convex in *all* unknowns, but the convexity with respect to u facilitates the computation of the transformation parameters.

Possible Priors. Points on the view sphere predicted by the motion model can be used to provide a prior when segmenting subsequent frames of an image sequence. This can be done in several ways — the most obvious is to take the shape at $p_{pred} \in \mathbb{S}^2$ from Equation (5.4) as a template. To incorporate the prior into the segmentation method, it is most appealing to impose

a vector field defined on a contour C that drives C along a geodesic in shape space towards the prior; this appears to be a sensible choice and has been proposed amongst others in [82]. Parametric active contour methods seem to be naturally suited for this sort of modification, since they work directly on points lying on the contour. For the implicit level set method [31, 110] or the method described in [33], applying a vector field that is defined only on the level set defining the interface is a little more involved. Imposing a flow along a geodesic in the implicit framework for other distance measures has been proposed, e.g., in [137]. The prior we use is a single shape predicted by the motion model on the view sphere. The shape is interpolated using a weighted Karcher mean and converted to a binary image. This binary image is then used as a prior for segmentation.

5.5 Experiments and Evaluation

Figures 5.5 and 5.6 show the results of the following experiment: for a given sequence $\{I_1, \ldots, I_n\}$ of images depicting a moving object, the contour c_1 and view sphere position t_1 for the first image were initialized manually. Then, using the methods from Sections 5.3 and 5.4, for each subsequent image I_{i+1} the contour c_{i+1} and the respective view sphere point t_{i+1} were updated. The contour c_i from the previous image was used for initialization and as a weak prior for the segmentation of image I_{i+1}. The segmentation result from I_{i+1} was then used to calculate t_{i+1}, starting at t_i, using the method described in Section 5.3.
In Fig. 5.6, an occluding object was added in a different scene, which could be successfully handled by using c_i as prior template for the segmentation algorithm. For these experiments, the level set algorithm was used. The figures depict a few snapshots from the whole sequences, which respectively consist of 100 and 50 frames each. These experiments show that the sphere tracking mechanism is capable of keeping track of the view sphere position fairly well, given a sufficient number of samples on the view sphere for interpolating the shape submanifold corresponding to the object. Fig. 5.7 shows results for a real recorded sequence.

5.6 Conclusions and Further Work

We presented a method that combines techniques from elastic shape manifold modeling, segmentation and optimization, to track the change of pose of a 3D object through tracking its contour. While the given contours of the object are currently sampled more or less uniformly

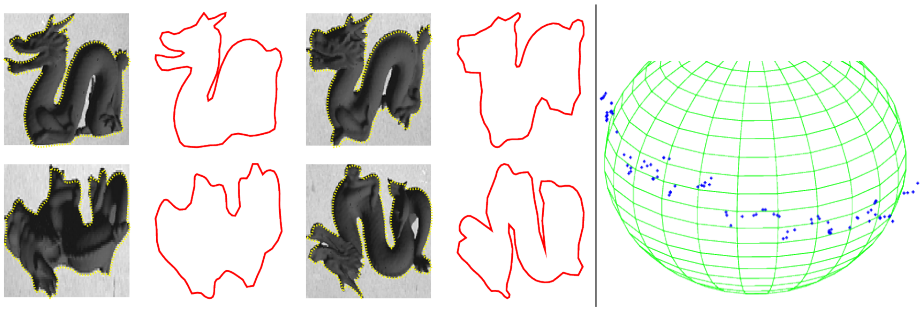


Figure 5.5: Experiment tracking the view sphere position using only the segmented contours from a sequence of images. Right: shown are measurements obtained on the view sphere, for the complete sequence. Left: a few images from the sequence are shown, the corresponding interpolated contours from the shape space $\mathcal{C}$ to their right. The initial position $t_0 \in \mathbb{S}^2$ and shape s_0 were given manually. Then for each image, the result from the previous one was used as initialization. A region based level set segmentation was used, with a curvature regularization term after [51].

on the view sphere, an adaptive sampling strategy may be investigated in future work: the amount of contour change depends on the position on $\mathbb{S}^2$ and the object in question. Advanced sampling should adapt the density of points in areas of rapid shape change on $\mathbb{S}^2$, thus exploiting the geometry of the shape submanifold already during data acquisition. However, in our experiments sampling 162 points appeared to be sufficient.

Another point concerns initialization, which is currently done manually. Automatic initialization may be achieved for example by a voting scheme on the first few frames, for sequences where the first few contours can be extracted well enough by any extraction method.

Regarding the segmentation prior, another option is to investigate weighted combination of a local neighborhood of shapes around p_{pred} to create a template with a "fuzzy" boundary, in order to take more into account inherent uncertainties of the predicted path of shapes.

A last matter worth mentioning is computation speed. Specifically, potential for speed-up is in the numerical calculation of the Log map for $\mathcal{C}$.

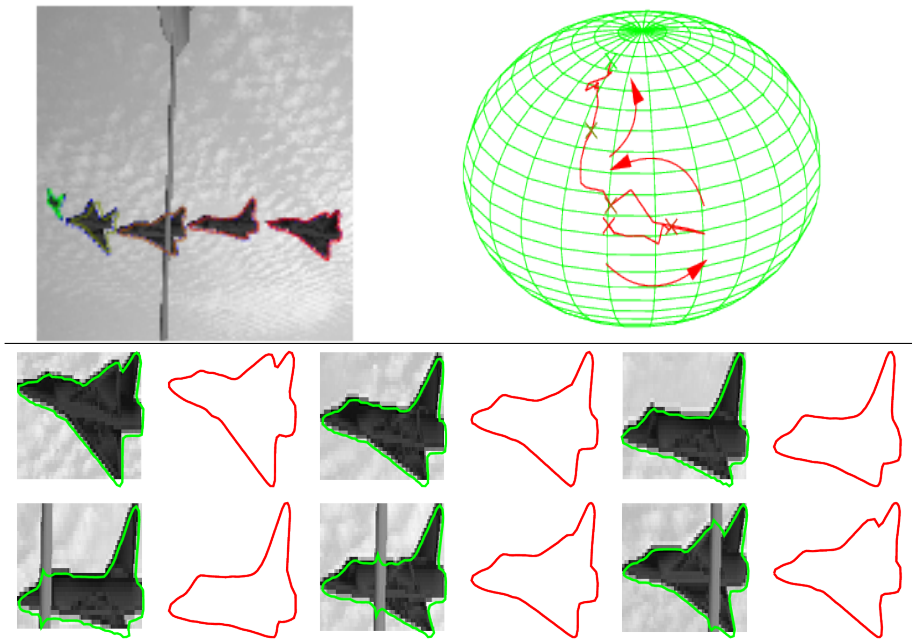


Figure 5.6: Sphere tracking experiment with occlusion. The top row shows the tracked view sphere path on the right (the arrows indicate the direction of motion), and an illustration of the image sequence on the left. The color coding shows the corresponding contours and view sphere positions. Using the resulting shape from each previous frame to create a prior for the segmentation algorithm enables the sphere tracking to keep going for this sequence, where a small occluding object moves in front of the object. Each row shows the area of interest from 3 subsequent frames with the superimposed segmentation result, followed by the contour representing the shape tracked on the view sphere.

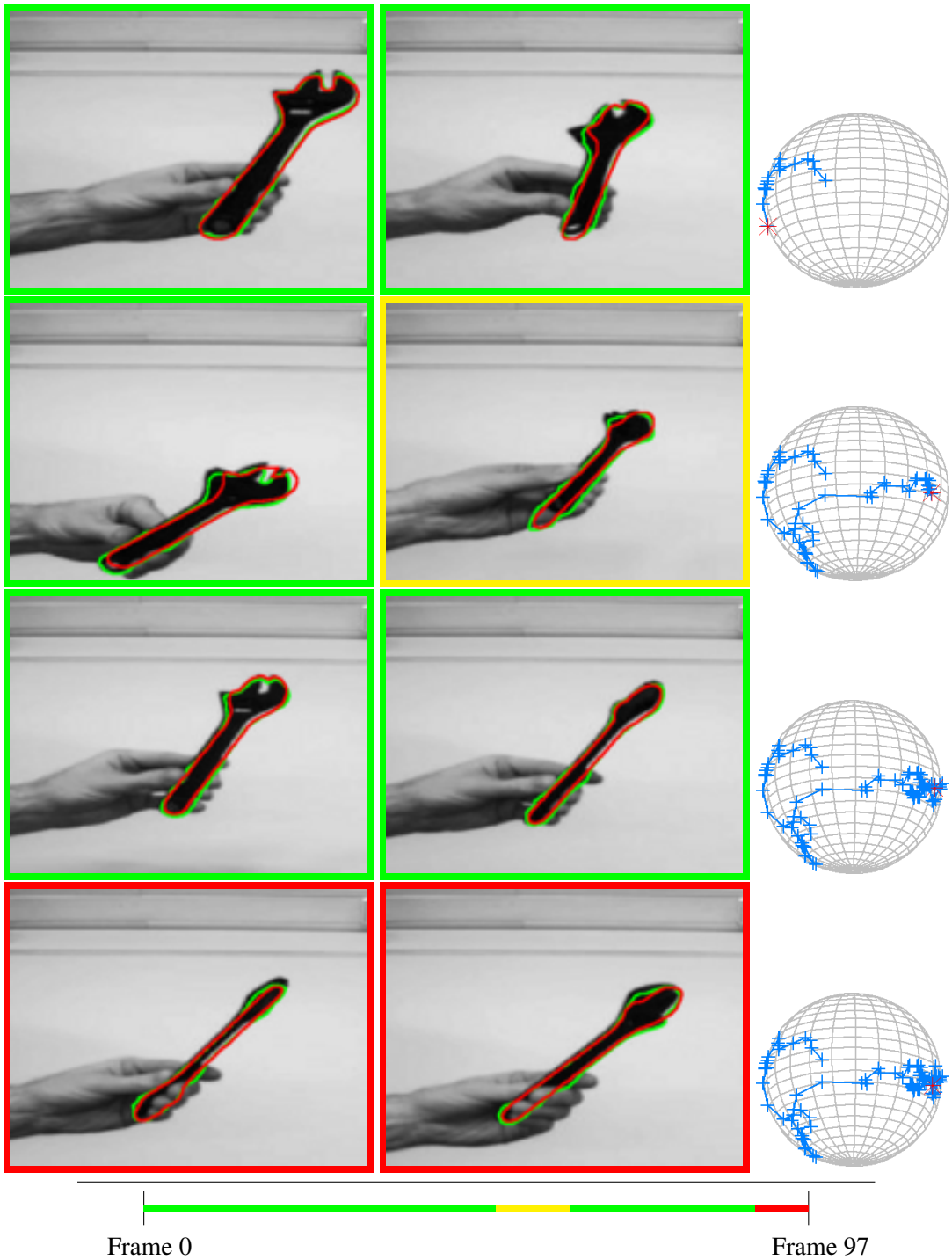


Figure 5.7: Sphere tracking with a real recorded sequence totalling 97 frames. Roughly every 20^{th} is shown. Indicated in each frame are the segmentation result (green) and aligned interpolated shape (red). Difficult situations where the view tracking goes wrong are indicated in red, yellow are situations which are just ok. The time line on the bottom indicates the situation for the whole 97 frames.The spheres on the right indicate the inferred view positions along the sequence.

5. VIEW POINT TRACKING OF RIGID OBJECTS USING REGION-BASED SEGMENTATION MODEL AND SHAPE PRIORS FROM SHAPE SUBMANIFOLDS

Chapter 6

Pose Invariant Shape Prior Segmentation using Continuous Cuts and Gradient Descent on Lie Groups

---

This chapter is a reformatted and slightly modified version of paper [112] without affecting the content or the results. Here we propose a novel formulation of the Chan-Vese model for pose invariant shape prior segmentation as a continuous cut problem. The model is based on the classic L^2 shape dissimilarity measure and with pose invariance under the full (Lie-) group of similarity transforms in the plane. To overcome the common numerical problems associated with step size control for translation, rotation and scaling in the discretization of the pose model, a new gradient descent procedure for the pose estimation is introduced. This procedure is based on the construction of a Riemannian structure on the group of transformations and a derivation of the corresponding pose energy gradient. Numerically, this amounts to an adaptive step size selection in the discretization of the gradient descent equations. Together with efficient numerics for TV-minimization we get a fast and reliable implementation of the model. Moreover, the theory introduced is generic and reliable enough for application to more general segmentation and shape models.

6.1 Introduction

The celebrated model of T. Chan and L. Vese [31] for piecewise constant, two-phase segmentation of a gray scale image $I : \Omega \to \mathbb{R}_+$ can be formulated as follows: Among all characteristic functions $u = \mathbf{1}_\Sigma$ of measurable sets Σ, contained in the bounded (image) domain $\Omega \subset \mathbb{R}^2$, and all pairs of real numbers $\mathbf{c} = (c_0, c_1)$, find $u^* = \mathbf{1}_{\Sigma^*}, \mathbf{c}^* = (c_0^*, c_1^*)$ which minimizes the following energy

$$E_{\mathrm{CV}}(u, \mathbf{c}) = J(u) + \frac{\lambda}{2}\Big\{ \langle 1-u, (I-c_0)^2 \rangle + \langle u, (I-c_1)^2 \rangle \Big\}, \tag{6.1}$$

where $\lambda > 0$ is a fixed weight, $J(u) = \int_\Omega |\nabla u|\, dx$ is the total variation of u, and $\langle u, v \rangle = \int_\Omega uv\, dx$ is the L^2 inner product between u and v. Recall that for $u = \mathbf{1}_\Sigma$, $J(u) = \mathrm{Per}(\Sigma)$, the perimeter (in Ω) of Σ, i.e. the length of the boundary $\Gamma = \partial\Sigma$ in Ω.

Traditionally, and originally [31], minimization of (6.1) was formulated in the level set framework of Osher an Sethian [110, 111, 134] by setting $u = H(\phi)$, where H denotes the Heaviside function, and $\phi : \Omega \to \mathbb{R}$ an embedding function used to represent the image object implicitly as $\Sigma = \{x \in \Omega\,;\, \phi(x) > 0\}$. This highly non-linear optimization problem is solved using gradient descent, which, in the level set framework, corresponds to the following evolution PDE for the *active contour* $\Gamma(t) := \partial\Sigma(t) = \{x \in \Omega\,;\, \phi(x,t) = 0\}$,

$$\frac{\partial \phi}{\partial t} = \mathrm{div}\Big(\frac{\nabla \phi}{|\nabla \phi|}\Big) + \frac{\lambda}{2}\left[(I-c_0)^2 - (I-c_1)^2\right] |\nabla \phi|,$$

where t is an artificial time parameter and $\phi = \phi(x,t)$ a time dependent level set function. At every instant in this evolution, the gray value estimates c_0, c_1 are updated according to

$$c_0 = c_0(u) = \frac{\langle 1-u, I \rangle}{\langle 1-u, 1 \rangle} \quad \text{and} \quad c_1 = c_1(u) = \frac{\langle u, I \rangle}{\langle u, 1 \rangle}. \tag{6.2}$$

One of the most inspiring discoveries in recent years, due to Chan, Esedoūglu and Nikolova [33], is that, for any fixed $\mathbf{c}$, the minimization (6.1) with respect to binary label functions u may be solved *exactly* by considering a convex relaxation of the problem, where the set of admissible u's is enlarged to:

$$K := \{u \in BV(\Omega)\,;\, 0 \le u(x) \le 1 \text{ for all } x \in \Omega\}. \tag{6.3}$$

In fact, it was shown in [33] that if $u \in K$ minimizes (6.1), then for almost all thresholds

$t \in (0,1)$ the function

$$u^t(x) = \begin{cases} 1 & \text{if } u(x) > t \\ 0 & \text{otherwise} \end{cases} \qquad (x \in \Omega), \tag{6.4}$$

is a global minimizer for the original problem. The proof is recalled in Section 6.2.1. Thus, global minimizers of the Chan-Vese model can be found by truncation of the solution to an easier, unilaterally constrained, convex variational problem. The use of this truncation property is referred to as the *continuous (graph) cut* method, and problems formulated in this manner can be solved efficiently using fast algorithms for TV-minimization. See, e.g., Chambolle [28].

The problem of including a priori shape information into the segmentation process has been studied extensively within the level set framework for the last decade or so [32, 42, 93, 127, 128]. The common approach is to include a interaction energy between object Σ and a prior shape Σ' into the segmentation functional. If f denotes the characteristic function of the prior shape Σ', then a typical shape prior segmentation functional looks like

$$E(u, \mathbf{c}, f) = E_{\mathrm{CV}}(u, \mathbf{c}) + \frac{\gamma}{2}\|u - f\|^2, \tag{6.5}$$

where $\gamma > 0$ is a fixed coupling constant for the interaction, and $\|u\| = \sqrt{\langle u, u \rangle}$ is the L^2 norm. The shape interaction in (6.5) may be interpreted geometrically as $\|u - f\|^2 = \mathrm{Area}(\Sigma \triangle \Sigma')$, i.e. the area of the symmetric set difference between the sets Σ and Σ', c.f. [32] and [127]. The segmentation is now obtained by minimization of the functional (6.5) with respect to the (binary) label functions u, gray values $\mathbf{c}$ and $f \in \mathcal{F}$, where $\mathcal{F}$ denotes a class of prescribed shape priors. This formulation is quite general. A specific example, considered in this chapter, is segmentation with pose invariant priors. In this case $\mathcal{F} = \{f = f_0 \circ T\}_{T \in G}$, where the binary function f_0 is a shape template, and T ranges over a group of transformations G, e.g. the group of similarity transformations.

Since continuous cuts have emerged as an alternative to level sets for minimization of the CV- and other segmentation models, it is natural to ask if known shape prior segmentation models can be reformulated as variational problems possessing the important truncation property, which allow them to be solved using TV-minimization algorithms. One such attempt has been made in [64], see Section 6.2.2, but it does not go all the way.

The purpose of the present work is to formulate the shape prior segmentation model (6.5) as a continuous cut problem. This is achieved by reformulating the problem as a CV model (see Section 6.3.1). We specifically consider shape priors which are pose invariant under the

group of similarity transforms, which involves optimization over a Lie group. In order to solve this problem efficiently and reliably, we develop a theory for gradient descent on Lie groups (Section 6.3.3). The problem here is, essentially, to construct a Riemannian structure on the Lie group. The new theory eliminates the problems associated with step-size selection in discretizations of the gradient descent ODEs usually encountered in segmentation models with pose estimation.

6.2 Background: Relaxation in the CV Model

6.2.1 Relaxation in the CV Model

In this section we briefly describe the theory behind the continuous cut solution for the CV model and its connection to the ROF denoising model and TV-minimization. Let us consider the minimization of (6.1) over the set of label functions $u \in K$ defined in (6.3), and gray values $\mathbf{c} \in \mathbb{R}^2$. In this setting E_{CV} is a bi-convex functional, that is, convex in each of its arguments u and $\mathbf{c}$, separately, when the other is kept fixed. However, E_{CV} is not jointly convex. One therefore uses a method referred to as the *CV-algorithm*, which alternates between optimization in u and $\mathbf{c}$: If an initial state $(u^0, \mathbf{c}^0)$ is given, then a minimizing sequence $(u^k, \mathbf{c}^k)$ is constructed by

$$u^{k+1} = \arg\min_{u \in K} E_{\text{CV}}(u, \mathbf{c}^k), \tag{6.6}$$

$$\mathbf{c}^{k+1} = \arg\min_{\mathbf{c} \in \mathbb{R}^2} E_{\text{CV}}(u^{k+1}, \mathbf{c}). \tag{6.7}$$

The sub-problem (6.7) is a simple quadratic optimization whose solution is readily given by the formulas in (6.2) with $u = u^{k+1}$. We therefore proceed to describe the theory and algorithms for the continuous cut solution of the sub-problem (6.6). If $\mathbf{c}$ is fixed then the minimization of (6.1) over K is equivalent to minimization over K of the energy

$$\hat{E}(u) = J(u) + \frac{\lambda}{2}\langle (I-c_1)^2 - (I-c_0)^2, u\rangle := J(u) + \langle g, u\rangle, \tag{6.8}$$

where $g = (\lambda/2)[(I-c_1)^2 - (I-c_0)^2]$ is the data term. We now prove the result in Chan *et al.* [33] referred to in the Introduction, that minimization of $\hat{E}$ over binary u's can be obtained from the solution of the convex variational problem $\inf_{u \in K} \hat{E}(u)$ by truncation. For $u \in BV(\Omega)$, let u^t denote the function defined in (6.4). We have the result

The Truncation Lemma *If $u \in K$ solves $\inf_K \hat{E}$, then so does u^t for almost all $t \in [0,1]$.*

Proof The coarea formula, $J(u) = \int_0^1 J(u^t)\,dt$, and the layer cake representation $\langle g,u\rangle = \int_0^1 \langle g,u^t\rangle\,dt$, together yield $\hat{E}(u) = \int_0^1 \hat{E}(u^t)\,dt$. Since $u^t \in K$ it is admissible, and $\hat{E}(u^t) \geq \hat{E}(u)$ for all t, by assumption, the integrand on the left hand side of $\int_0^1 \hat{E}(u^t) - \hat{E}(u)\,dt = 0$ must be zero for almost all $t \in [0,1]$. ∎

In Chan *et al.* [33], the minimum was approximated by solving a degenerate parabolic PDE for u (the gradient descent PDE) with an exact penalty term to ensure that the constraint $0 \leq u \leq 1$ is satisfied at all times. This PDE was implemented with an explicit finite difference scheme, and is therefore rather slow. We have chosen another method, introduced by Aujol and Chambolle [63] and used successfully by Bresson *et al.* [20]. This consists of minimizing a variant of (6.8) which has been regularized slightly by infimal convolution with a quadratic function:

$$\inf_{v\in BV,\, u\in K} \big\{ J(v) + \frac{1}{2\theta}\|v-u\|^2 + \langle g,u\rangle \big\}, \tag{6.9}$$

where $\theta > 0$ is a parameter. This problem is solved iteratively using what we call the *ABC-algorithm*: If (v^0,u^0) denotes an initial guess, then a minimizing sequence is given by the pair (v^n,u^n) where

$$v^{n+1} = \operatorname*{arg\,min}_{v\in BV} \big\{ J(v) + \frac{1}{2\theta}\|v-u^n\|^2 \big\} = u^n - \theta \Pr_C(u^n/\theta), \tag{6.10}$$

$$u^{n+1} = \operatorname*{arg\,min}_{u\in K} \big\{ \frac{1}{2\theta}\|v^{n+1}-u\|^2 + \langle g,u\rangle \big\} = \Pr_K(v^n - \theta g). \tag{6.11}$$

The first of these problems is the classical Rudin-Osher-Fatemi (ROF) image denoising model [132] with u^n as input image. The second one is a simple L^2-optimization. Both problems are strictly convex, thus admits unique solutions, and, as indicated, their optima can be expressed in terms of L^2-projections onto closed convex sets: the first projection is onto C, which is the L^2-closure of the set $\{\operatorname{div}\xi\,;\,\xi \in C^1(\Omega;\mathbb{R}^2), |\xi(x)| \leq 1\,\forall x \in \Omega\}$, c.f. Chambolle [28]. The second projection is onto K, defined above. The latter is easy to compute, indeed $\Pr_K f(x) = \min(1,\max(0,f(x)))$ for $x \in \Omega$), for any square L^2 function $f : \Omega \to \mathbb{R}$. To minimize the ROF functional (6.10) we use a variant of the fast and reliable algorithm for TV-minimization proposed by Chambolle [28, 29].

6.2.2 Continuous Cuts for Prior-Based Object Segmentation

Chapter (4) contains an attempt to include shape priors into continuous cut segmentation that considers the model (6.5) where $f = f_0 \circ T$ is pose invariant under the group of similarity transformations T of the plane, i.e. the variational problem

$$\inf_{u,\mathbf{c},T} \left\{ E(u,\mathbf{c},T) := E_{\mathrm{CV}}(u,\mathbf{c}) + \frac{\gamma}{2}\|u - f_0 \circ T\|^2 \right\}. \tag{6.12}$$

This problem cannot be solved by continuous cuts (for $\mathbf{c}$ and T fixed) simply by enlarging the admissible label functions from the binary u's to $u \in K$. The problem, of course, lies in the quadratic interaction term, which seems to "spoil" the Truncation Lemma. As proposed in [64] this problem is circumvented by the following construction: If $(u^0, \mathbf{c}^0, T^0)$ denotes an initial guess then a minimizing sequence $(u^k, \mathbf{c}^k, T^k)$ is (essentially) constructed by the following procedure:

$$\mathbf{c}^{k+1} = \mathbf{c}(u^k) \quad \text{using formula (6.2).} \tag{6.13}$$

$$T^{k+1} = T^k - \Delta t \frac{\partial}{\partial T} E(u^k, \mathbf{c}^{k+1}, T^k) \quad \text{time step } \Delta t > 0 \tag{6.14}$$

$$u^{k+1} = \arg\min_{u \in K} E_{\mathrm{CV}}(u, \mathbf{c}^{k+1}) + \frac{\gamma}{2}\langle u - f_0 \circ T^{k+1}, u^k - f_0 \circ T^{k+1} \rangle \tag{6.15}$$

Here we observe that by freezing one occurrence of $u = u^k$ in the quadratic interaction term, update step (6.15) becomes linear in u, hence solvable by continuous cut methods. In [64] this minimization was performed using the gradient descent PDE from [33]. Our aim is to improve the above method by formulating the problem in such a way that the model itself, not only the algorithm, satisfies the truncation property.

6.3 The Shape Prior Segmentation Model

6.3.1 The Basic Energy Functional

Our reformulation of the functional (6.5) is based on the following observation: If the label function $u : \Omega \to \{0,1\}$ is binary, and we define an image model by $I_{\mathrm{model}} = I_{\mathrm{model}}(u,\mathbf{c}) = c_0(1-u) + c_1 u$, then it is easy to see that the CV-functional (6.1) may be rewritten as:

$$E_{\mathrm{CV}}(u,\mathbf{c}) = J(u) + \frac{\lambda}{2}\|I - I_{\mathrm{model}}\|^2. \tag{6.16}$$

This suggests the following model for shape prior segmentation: If $f : \Omega \to \mathbb{R}$ denotes a (possibly fuzzy) shape prior, that is $0 \leq f(x) \leq 1$ on Ω, then we associate an image model to f given by $I_{\text{prior}} = I_{\text{prior}}(f, \mathbf{b}) = b_0(1-f) + b_1 f$. We now pose shape prior segmentation as the minimization over all binary label functions u of the following functional:

$$E(u, \mathbf{c}, f, \mathbf{b}) = E_{CV} + E_{\text{prior}} = J(u) + \frac{\lambda}{2}\|I - I_{\text{model}}\|^2 + \frac{\mu}{2}\|I_{\text{model}} - I_{\text{prior}}\|^2. \tag{6.17}$$

Notice that close to convergence, it is reasonable to expect that $b_0 \approx c_0$ and $b_1 \approx c_1$. Assuming that exact equality holds here, we find that

$$\frac{\mu}{2}\|I_{\text{model}} - I_{\text{prior}}\|^2 = \frac{\mu}{2}(c_1 - c_0)^2\|u - f\|^2, \tag{6.18}$$

which corresponds to the interaction term in (6.5) if we set $\gamma = \mu(c_1 - c_0)^2$. We will use this simplification in Section 6.3.2.

Let us consider the minimization of (6.17) with respect to u and $\mathbf{c}$ when prior data $\mathbf{b}$ and f are kept fixed. After completion of squares in (6.17) we find that

$$\begin{aligned} E(u, \mathbf{c}, f, \mathbf{b}) = J(u) &+ \frac{\lambda+\mu}{2}\Big\|I_{\text{model}} - \Big(\frac{\lambda}{\lambda+\mu}I + \frac{\mu}{\lambda+\mu}I_{\text{prior}}\Big)\Big\|^2 \\ &+ \frac{\lambda+\mu}{2}\Big\{\frac{\lambda}{\lambda+\mu}\|I\|^2 + \frac{\mu}{\lambda+\mu}\|I_{\text{prior}}\|^2 - \Big\|\frac{\lambda}{\lambda+\mu}I + \frac{\mu}{\lambda+\mu}I_{\text{prior}}\Big\|^2\Big\}. \end{aligned} \tag{6.19}$$

Only the first square depends on the (binary) u and $\mathbf{c}$. So updating u and $\mathbf{c}$ is equivalent to solving the following CV-problem:

$$\inf\Big\{J(u) + \frac{\lambda+\mu}{2}\big[\langle 1-u, (I_{\text{eff}} - c_0)^2\rangle + \langle u, (I_{\text{eff}} - c_1)^2\rangle\big]\Big\}. \tag{6.20}$$

Here $I_{\text{eff}} = \frac{\lambda}{\lambda+\mu}I + \frac{\mu}{\lambda+\mu}I_{\text{prior}}$ is an effective image obtained as a convex combination of the observed image I and the prior image I_{prior}. The problem (6.20) has the truncation property, and may be solved by the CV-algorithm (6.6), (6.7), using continuous cuts. This solution is a minimizer of (6.17).

Suppose that $\mathbf{c}$ and u have been updated and are now held fixed. Returning to the energy E, written in the original form (6.17), we optimize with respect to prior image model $I_{\text{prior}} = b_0(1-f) + b_1 f$. An easy calculation shows the optimal gray scales $\mathbf{b} = (b_0, b_1)$ are given by

the formulas:

$$b_0 = \frac{\langle 1-f, I_{\text{model}} \rangle}{\|1-f\|^2} \quad \text{and} \quad b_1 = \frac{\langle f, I_{\text{model}} \rangle}{\|f\|^2}.$$

With these values fixed, we proceed to update the pose of the shape prior f, which is the subject of the next few sections.

6.3.2 Pose Invariant Prior Interaction Energy

Let $f_0 : \Omega \to \mathbb{R}$ denote a shape template of class $C_0^1(\Omega)$, and $T : \mathbb{R}^2 \to \mathbb{R}^2$ a similarity transformation, that is, a mapping of the form $y = T(x) = \mu^{-1}R^{-1}(x-\mathbf{a}), \quad x \in \mathbb{R}^2$, where $R \in \mathrm{SO}(2)$ denotes rotation, $\mu > 0$ a scaling factor, and $\mathbf{a} \in \mathbb{R}^2$ translation. We define the shape prior f as the transformed template $T^* f_0 : \mathbb{R}^2 \to \mathbb{R}$ by the formula $f(x) = T^* f_0(x) = (f_0 \circ T)(x) = f_0(T(x))$ for all $x \in \mathbb{R}^2$. If T is sufficiently close to the identity map then, clearly, $T^* f_0 \in C_0^1(\Omega)$, so that the support of the prior will remains inside the image domain Ω.

In this chapter we use the simplification of (6.17) in (6.18) and consider a pose invariant prior interaction defined by the energy,

$$E_{\text{prior}}(u) = \inf_T \|u - T^* f_0\|^2 = \inf_T \int_\Omega (u(x) - f_0(T(x)))^2 \, dx, \tag{6.21}$$

where the infimum is taken over the group of similarity transforms T in the plane. The following (natural) parametrization is used throughout:

$$\mathbf{a} \in \mathbb{R}^2, \quad \mu = \mathrm{e}^\sigma \ (\sigma \in \mathbb{R}), \text{ and } R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \ (\theta \in \mathbb{R}). \tag{6.22}$$

The pose parameters θ, σ and $\mathbf{a}$ are collected in a vector $\mathbf{p} = (p_1, p_2, p_3, p_4) := (\theta, \sigma, \mathbf{a}) \in \mathbb{R}^4$, the corresponding map is occasionally denoted $T = T(\mathbf{p})$, and the shape prior becomes $f(x) = T^* f_0(x) = T(\mathbf{p})^* f_0(x) = f_0(\mathrm{e}^{-\sigma} R(-\theta)(x-\mathbf{a}))$.

Now, the infimum in (6.21) is usually computed by applying a gradient descent procedure to the function $\mathbb{R}^4 \ni \mathbf{p} \mapsto E(\mathbf{p}) := \|u - T(\mathbf{p})^* f_0\|^2/2$. That is, one solves a system of ODE:s given by $\mathbf{p}'(t) = -\nabla E(\mathbf{p}(t))$, with respect to an artificial time parameter t, and the obtain the optimal pose $\mathbf{p}^*$ as $\mathbf{p}^* = \lim_{t\to\infty} \mathbf{p}(t)$. This method requires the computation of the partial derivatives $\partial E(\mathbf{p})/\partial p_i$ for every component p_i of $\mathbf{p}$. A simple calculation shows that $\partial E(\mathbf{p})/\partial p_i = \langle T(\mathbf{p})^* f_0 - u, \partial T(\mathbf{p})^* f_0/\partial p_i \rangle$, so we begin with the partials $\partial T(\mathbf{p})^* f_0(x)/\partial p_i$.

By the chain rule,

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{a}} T^* f_0(x) &= -\nabla_x T^* f_0(x) = -\nabla_x f(x) \quad \text{(two components!)} \\
\frac{\partial}{\partial \theta} T^* f_0(x) &= -\nabla_x T^* f_0(x)^T J(x-\mathbf{a}) = -\nabla_x f(x)^T J(x-\mathbf{a}) \\
\frac{\partial}{\partial \sigma} T^* f_0(x) &= -\nabla_x T^* f_0(x)^T (x-\mathbf{a}) = -\nabla_x f(x)^T (x-\mathbf{a})
\end{aligned}$$

where $J = R(-\theta)^T R'(-\theta) = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$ is the clockwise rotation by $\pi/2$ radians. Notice that $-\nabla_x f$ appears in all the formulas, with the x-derivative computed after transformation of the template.

It follows from the above formulas that the partial derivatives of $E(\mathbf{p})$ are given by

$$\begin{gathered}
\frac{\partial}{\partial \mathbf{a}} E(\theta,\sigma,\mathbf{a}) = -\langle f-u, \nabla_x f\rangle, \quad \frac{\partial}{\partial \theta} E(\theta,\sigma,\mathbf{a}) = -\langle f-u, \nabla_x f^T J(\cdot-\mathbf{a})\rangle, \\
\text{and} \quad \frac{\partial}{\partial \sigma} E(\theta,\sigma,\mathbf{a}) = -\langle f-u, \nabla_x f^T (\cdot-\mathbf{a})\rangle.
\end{gathered} \tag{6.23}$$

These integrals are effectively computed on the support of $-\nabla_x f$, that is, over a neighborhood of the boundary of the shape prior.

The traditional way to proceed is to iteratively update the pose parameters $\mathbf{a}, \theta$ and σ using (essentially) the schemes $\mathbf{a}(t+\Delta t_{\mathbf{a}}) = \mathbf{a}(t) - \Delta t_{\mathbf{a}} \cdot \partial E/\partial \mathbf{a}$, $\theta(t+\Delta t_\theta) = \theta(t) - \Delta t_\theta \cdot \partial E/\partial \theta$, and $\sigma(t+\Delta t_\sigma) = \sigma(t) - \Delta t_\sigma \cdot \partial E/\partial \sigma$. This is problematic; in order for this method to work properly the time steps $\Delta t_{\mathbf{a}}, \Delta t_\theta$ and Δt_σ have to be chosen differently, and with great care. This is not only unsatisfying from a theoretical view point but it also limits the practical applicability of the method; not least because the delicate choice of time steps tends to be time-consuming. We address this problem in the next section.

6.3.3 The Gradient Construction

The group of similarity transformations constitutes a four-dimensional manifold that we denote $\mathcal{M}$ (i.e., $\mathcal{M}$ is a Lie group). Any point $p \in \mathcal{M}$ may be represented by the coordinates $\mathbf{p} = (\theta, \sigma, \mathbf{a})$ using (6.22), which may be regarded as a (almost global) parametrization of a neighbourhood of the identity map in $\mathcal{M}$. If $E : \mathcal{M} \to \mathbb{R}$ is a differentiable function then $dE(p) : T_p\mathcal{M} \to \mathbb{R}$ denotes the differential of E at $p \in \mathcal{M}$, where $T_p\mathcal{M}$ is the tangent space of $\mathcal{M}$ at p. In the local coordinates the differential may be expressed as $dE = \frac{\partial E}{\partial \mathbf{a}} d\mathbf{a} + \frac{\partial E}{\partial \theta} d\theta + \frac{\partial E}{\partial \sigma} d\sigma$. Suppose that $T_p\mathcal{M}$ is equipped with a scalar product $(\cdot,\cdot)_p$, then we may define the gradient of

E at p as the unique vector $\nabla E(p) \in T_p\mathcal{M}$ which satisfies the relation

$$(\nabla E(p), v)_p = dE(p)v, \quad \forall v \in T_p\mathcal{M}. \tag{6.24}$$

The metric $ds^2 = |d\mathbf{a}|^2 + d\theta^2 + d\sigma^2$ defines a scalar product which, as already noted, is insufficient for the construction of a reliable gradient descent scheme for $E(\mathbf{p}) = \|u - T(\mathbf{p})^* f_0\|^2/2$. Our goal is to define a Riemannian structure on $\mathcal{M}$ which is better suited for this task.

Let a function $f : \mathcal{M} \times \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(\mathbf{p},x) = T(\mathbf{p})^* f(x) = f_0(T(\mathbf{p})x)$. Since the shape template $f_0 \in L^2(\mathbb{R}^2)$, the mapping $\mathbf{p} \mapsto f(\mathbf{p},\cdot)$ is a function $f : \mathcal{M} \to L^2(\mathbb{R}^2)$. Now, $L^2(\mathbb{R}^2)$ comes with an inner product $\langle\cdot,\cdot\rangle$, so it is natural to define the scalar product on $T_p\mathcal{M}$ as the *pullback* by f of the L^2-inner product to the tangent space $T_p\mathcal{M}$,

$$(v,w)_p = \langle df(p)v, df(p)w\rangle, \quad (v,w \in T_p\mathcal{M}) \tag{6.25}$$

where $df(p) : T_p\mathcal{M} \to T_{f(p)}L^2(\mathbb{R}^2) \equiv L^2(\mathbb{R}^2)$ denotes the differential of f. By the chain rule, $df = -(D_x f_0 \circ T)dT$, so in view of the identity $dT = DT d\mathbf{p} = D_x T(\mathbf{p})DT(0)d\mathbf{p}$ (which uses the group structure of $\mathcal{M}$) we see that $df = -\nabla_x f^T DT(0)\, d\mathbf{p}$, where $DT(0)$ is the linear map given by the block matrix: $DT(0) = \begin{bmatrix} I_{2\times 2} & J(x-\mathbf{a}) & (x-\mathbf{a}) \end{bmatrix}$.

As before, $J = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$. With this calculation we find that

$$\begin{aligned} \langle df(p)v, df(p)w\rangle &= \langle -\nabla_x f^T DT(0)\, d\mathbf{p}(v), -\nabla_x f^T DT(0)\, d\mathbf{p}(w)\rangle \\ &= d\mathbf{p}(v)^T \langle 1, DT(0)^T \nabla_x f \nabla_x f^T DT(0)\rangle d\mathbf{p}(w) := d\mathbf{p}(v)^T G(\mathbf{p}) d\mathbf{p}(w), \end{aligned}$$

where $G(\mathbf{p})$ denotes the *metric tensor* on $T_p\mathcal{M}$ expressed in the coordinates $\mathbf{p}$. If we define $M = \nabla_x f \nabla_x f^T$ then $G(\mathbf{p}) = \langle 1, g(\mathbf{p},\cdot)\rangle$ where $g(\mathbf{p},\cdot) : \mathbb{R}^2 \to \mathbb{R}^{4\times 4}$ is given by $g(\mathbf{p},x) = DT(0)^T M DT(0)$, which equals

$$\begin{bmatrix} M & MJ(x-\mathbf{a}) & M(x-\mathbf{a}) \\ (x-\mathbf{a})^T J^T M & (x-\mathbf{a})^T J^T MJ(x-\mathbf{a}) & (x-\mathbf{a})^T JM(x-\mathbf{a}) \\ (x-\mathbf{a})^T M & (x-\mathbf{a})^T MJ(x-\mathbf{a}) & (x-\mathbf{a})^T M(x-\mathbf{a}) \end{bmatrix}$$

This expression is, unfortunately, too complicated for our present purpose, so we need to make some simplification. This is achieved by approximating the structure tensor M by the simpler tensor $\frac{1}{2}|\nabla_x f|^2 I_{2\times 2}$. (There are some compelling reasons for doing so! For instance $g_{3,3} +$

$g_{4,4} = |\nabla_x f|^2 |x-\mathbf{a}|^2$.) With this simplification we get

$$g(\mathbf{p},x) = |\nabla_x f|^2 \begin{bmatrix} I_{2\times 2} & J(x-\mathbf{a}) & (x-\mathbf{a}) \\ (x-\mathbf{a})^T J^T & |x-\mathbf{a}|^2 & (x-\mathbf{a})^T J(x-\mathbf{a}) \\ (x-\mathbf{a})^T & (x-\mathbf{a})^T J(x-\mathbf{a}) & |x-\mathbf{a}|^2 \end{bmatrix},$$

where we notice that, in fact, the matrix elements $g_{4,3} = g_{3,4} = 0$ because J is skew-symmetric. Finally, if we choose $\mathbf{a}$—the center of rotation and scaling—such that $\langle |\nabla_x f|^2, x-\mathbf{a}\rangle = 0$, that is, as the barycenter of the mass-distribution $dm = |\nabla_x f|^2\, dx$, then the metric tensor $G = \langle 1, g\rangle$ has the following diagonal form:

$$G(\mathbf{p}) = \begin{bmatrix} \|\nabla_x f\|^2 I_{2\times 2} & 0 & 0 \\ 0 & \| \, |x-\mathbf{a}| \nabla_x f \, \|^2 & 0 \\ 0 & 0 & \| \, |x-\mathbf{a}| \nabla_x f \, \|^2 \end{bmatrix}. \tag{6.26}$$

Equivalently, $(d\mathbf{p}, d\mathbf{p})_{\mathbf{p}} = \|\nabla_x f\|^2 \, |d\mathbf{a}|^2 + \| \, |x-\mathbf{a}| \nabla_x f \, \|^2 (d\theta^2 + d\sigma^2)$. It follows from (6.25) and the formulas (6.23), that the corresponding gradient of E has the components:

$$\begin{gathered} \nabla_{\mathbf{a}} E = \frac{\langle f-u, -\nabla_x f \rangle}{\|\nabla_x f\|^2}, \quad \nabla_\theta E = \frac{\langle f-u, -\nabla_x f^T J(\cdot - \mathbf{a}) \rangle}{\| \, |x-\mathbf{a}| \nabla_x f \, \|^2}, \\ \text{and} \quad \nabla_\sigma E = \frac{\langle f-u, -\nabla_x f^T (\cdot - \mathbf{a}) \rangle}{\| \, |x-\mathbf{a}| \nabla_x f \, \|^2}. \end{gathered} \tag{6.27}$$

This is the gradient used in our implementation of gradient descent search for the optimal pose parameters. Its use amounts to an adaptive step-size control in the numerical discretization of the associated system of ODEs.

6.4 Experiments

The method presented in Section 6.3 was implemented in MATLAB with the following specifics: For the minimization of (6.17) (in the form (6.19)) we used the ABC-algorithm (6.10) and (6.11) with the parameter $\theta = 0.5$ and a variant of Chambolle's algorithm [29, Eq. (12)], implemented with periodic boundary conditions, for the TV-minimization in (6.10). This was alternated with an update of the pose of the prior, using gradient descent with the new gradient (6.27).

The experiments presented here are limited to a proof-of-concept level. The first experiment (Figure 6.1) shows the CV segmentation with and without the shape prior, and with added

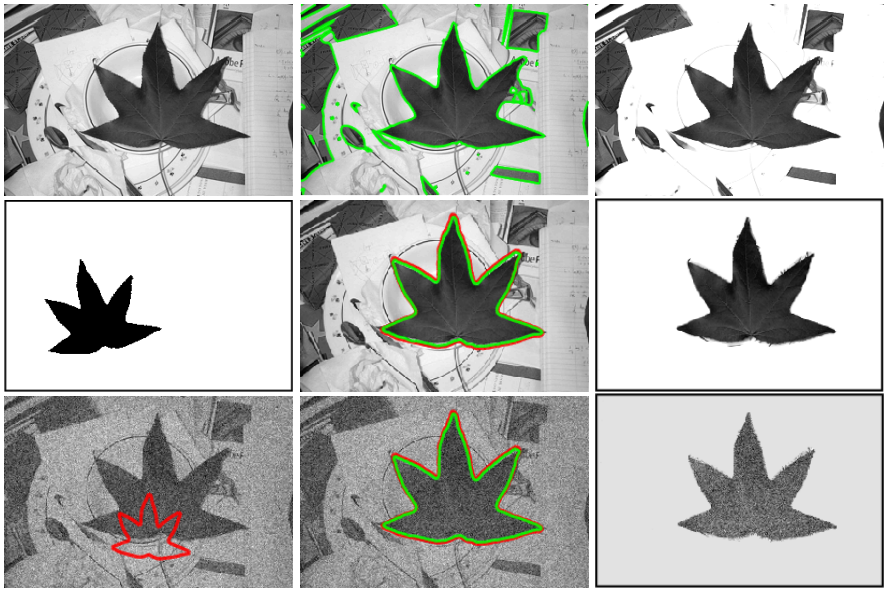


Figure 6.1: Experiment 1: First row: The original image, 212×320 pixels (left), the active contour $\Gamma = \{x; u(x) = .5\}$ in CV-segmentation without priors after 100 iterations (middle), and the corresponding segmentation (right). Second row: The shape template, the active contour and the shape prior after 150 iterations, and the final segmentation. Final row: segmentation of the image contaminated with 15% Gaussian noise using 200 iterations. Parameters: $\mu = .4$, $\lambda = .1$, $\theta = .5$ and step-size $\Delta t = .75$.

noise. The segmentation result is displayed as a cut-out from the original image by multiplication with the optimal label function u. This verifies the binary character of u. The second experiment (Figure 6.2) shows how the search evolves for three different initializations. As shown, the method may not always converge to the wanted solution. In fact, the prior contour may sometimes even shrink and disappear. These cases correspond, however, to quite plausible local minima for the pose energy, and this behavior is not unexpected in a local optimization method. More details are found in the figure captions.

6.5 Conclusions

This chapter contains two central contributions. Firstly, the reformulation in (6.17) of the shape prior segmentation model in (6.5), which leads to a minimization problem which can be solved using continuous cut methods. Secondly, the derivation of the gradient expressions (6.27), which is the basis for a stable and efficient gradient descent scheme for prior pose optimization. We believe that the ideas introduced here can be extended to cover more general and complex

shape prior segmentation models. In particular it would be interesting to see if the ideas can be applied to pose problems in three dimensions.

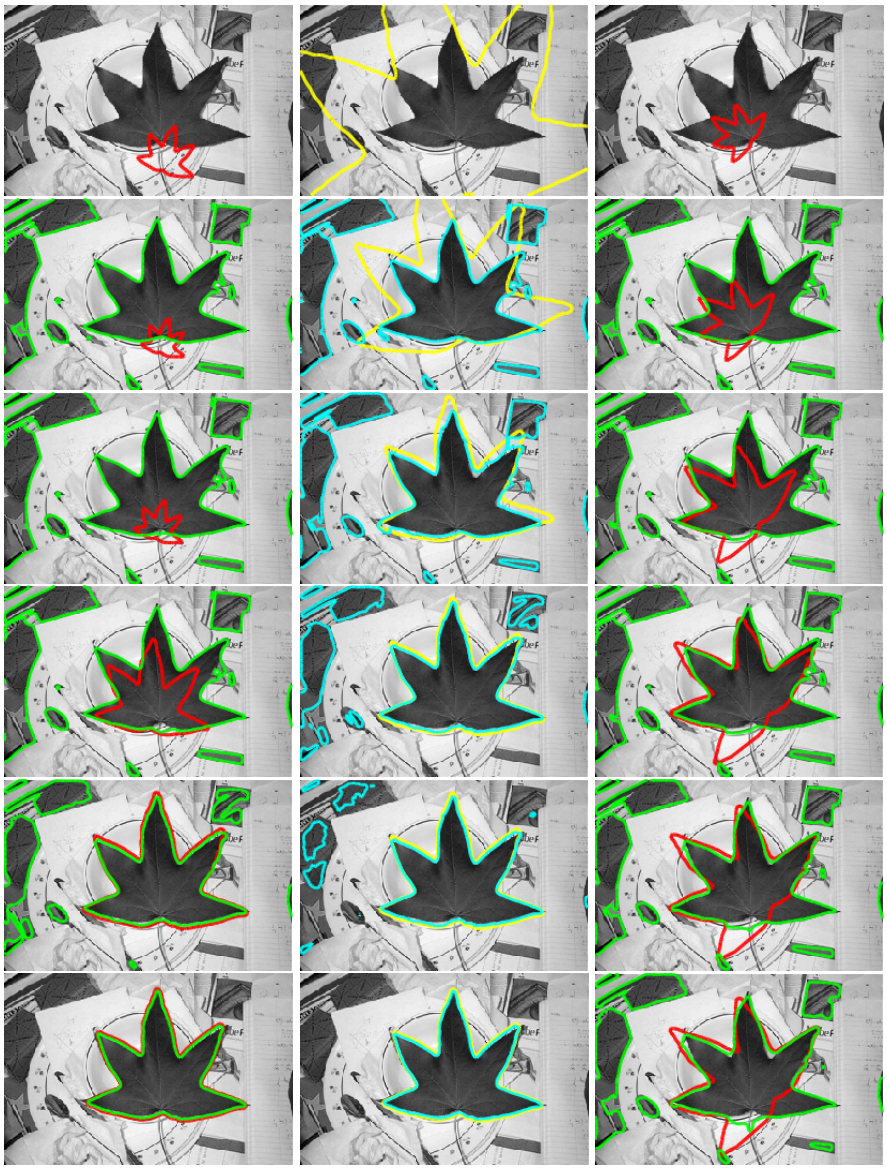


Figure 6.2: Experiment 2: Shape prior segmentation with three different initial poses (top row). Evolution after (approximately) 12, 25, 50, 100 and 200 iterations (rows 2–6). The run-time for 100 iterations is about 25 CPU-seconds. In the final phase of the segmentation, object previously detected outside the prior disappears. With the third initialization the he shape prior gets stuck in a local minimum. Such behavior cannot be ruled out when we work with local optimization methods. Image size and parameter settings are as in Experiment 1.

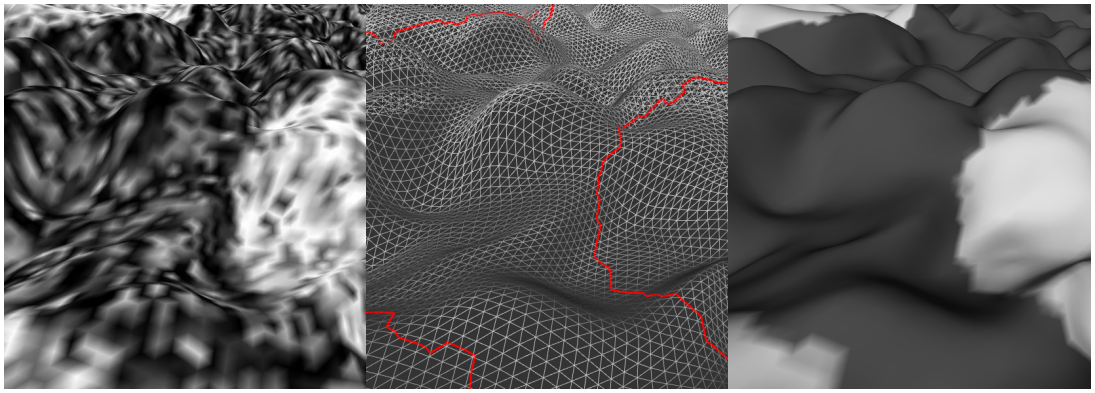
Chapter 7

Convex Multi-Region Segmentation on Manifolds

This chapter is a reformatted and slightly modified version of paper [48] without affecting the content or the results. Here we address the problem of segmenting data defined on a manifold into a set of regions with uniform properties. In particular, we propose a numerical method when the manifold is represented by a triangular mesh. Based on recent image segmentation models, our method minimizes a convex energy and then enjoys significant favorable properties: it is robust to initialization and avoid the problem of the existence of local minima present in many variational models. The contribution of this chapter is threefold: firstly we adapt the convex image labeling model to manifolds; in particular the total variation formulation. Secondly we show how to implement the proposed method on triangular meshes, and finally we show how to use and combine the method in other computer vision problems, such as 3D reconstruction. We demonstrate the efficiency of our method by testing it on various data.

7.1 Introduction

Image segmentation aims to partition a given image into several meaningful regions based on certain attributes such as intensity, texture, color, etc. This problem is one of the most challenging and important problems in computer vision. We address the problem of segmenting *data* defined on *manifolds* (typically a 2-surface in $\mathbb{R}^3$) into multiple regions of piecewise constant attributes. The ability to solve such a problem offers significant new possibilities in a number



(a) (b) (c)

Figure 7.1: Example of segmentation on a manifold. (a) The Input textured surface. (b) The triangular representation of the surface with the retrieved contour (in red). (c) Surface colored with the mean values of the segmented regions (and surface shading).

of applications. For example, in 3D reconstruction (see Jin et al. [78]), a segmentation into piecewise constant data of the reconstructed surface allows to naturally introduce constraints on the material of the scene.

7.1.1 Global Multi-Region Segmentation

Many approaches have been proposed to solve image segmentation problems. In particular, via gradient descents, variational methods for image segmentation have had a great success, such as snakes [83], geodesic active contours [27], geodesic active region [116] and the Chan-Vese models [31]. Yet, the main drawback of those methods is the existence of local minima due to the non-convexity of the energy functionals. Minimizing those functionals by gradient descent methods makes the initialization critical.

To obtain global minima, some previous image segmentation works have used different optimization techniques: For example the graph-cuts in a fully discrete setting, see [18, 86, 87] and the references therein. Nevertheless, while binary segmentation methods based on graph-cuts assure to get a global minima, multi-region segmentation algorithms are based on sequences of graph-cuts which cannot guarantee a global optimization.

Recently, some authors have tried to handle the problem in another direction. Instead of working on the optimization techniques in order to compute the minima of non-convex problems, they have reformulated the energy in order to get a fully convex problem [4, 20,

30, 33, 91, 121, 152]. These segmentation techniques are based on TV-regularizers and aim at finding characteristic functions that minimize the objective functions. Obtaining *global minima* becomes easy and can be done by simply performing a gradient descent. Also the *initialization problem vanishes*: the algorithm can start from any initialization and obtains the same result. The multi-region segmentation models proposed by [30, 91, 152] are rather similar, the work of Pock *et al.* [121] (inspired form Ishikawa's [76]) differs from the fact that it deals with ordered labels and uses a regularization term which favors transitions between nearby labels. This makes sense in their stereo application where the ordering is due to depth, but it is not the case when we deal with independent labels. Here, we adopt the model of [30, 91, 152] which is more appropriate to the applications we have in mind (the fact that real-life scenes are made by a finite number of independent materials, in 3D reconstruction problems) and we adapt this *image* labeling model to *manifolds*.

7.1.2 Data Segmentation on Manifolds

Manifolds such as surfaces are common in computer graphics as well as computer vision. Although data segmentation on surfaces has been recently used on implicit representations (see for example [78, 89]), explicit representations such as triangular meshes are natural and intuitive representations. Mesh representations have been widely used in 3D reconstruction, see for example [1, 49, 140], and recently in [1, 124, 153], which takes advantage of various recent evolution methods. These allow us to naturally deal with topological changes (necessary property e.g. in 3D shape estimation problems formulated within a variational framework). In a number of applications, particularly in graphics, this is the only representation one has at ones disposal. In this work, after adapting the image segmentation model to manifolds, we show how to implement the proposed method on *triangular meshes*.

Mesh segmentation has been used in computer graphics to decompose meshes into significant parts, but previous work mainly focus on the geometric aspects, the choice and the representations of features to use (like the curvature). We refer to [5] for a recent survey of those techniques, as in this work we focus in segmenting data on the mesh.

However, the problem of segmenting data like texture on manifolds has not received much attention until now, and is quite different from the geometric decomposition of a mesh. In [119], authors take into account both texture information and curvature, but their approach is based on a fast marching algorithm, which needs to be initialized using initial points. Moreover as their segmentation method is not convex, different regions sharing the same properties may

result in different labels. The same problem also occurs in [113], where watershed filtering was used. Contrary to those methods, the approach we propose is global and robust to initialization.

Finally, let us note that the segmentation model we consider is based on a total variation regularization. Although this regularization has been previously used on implicit surfaces (see [12] with applications to texture synthesis), it has not received much attention in Lagrangian methods. In the finite elements literature, one can find some papers dealing with the Laplace-Beltrami operator $\nabla_S \cdot \nabla_S u$ which corresponds to the gradient of the squared regularization term $\int_S |\nabla_S u|^2 d\sigma$ (see for example [52, 55]). To our knowledge there does not exist work dealing with the gradient of the total variation in this framework. Let us note here that the Laplace-Beltrami operator is linear while the term associated with the gradient of the total variation $\nabla_S \cdot \frac{\nabla_S u}{|\nabla_S u|}$ is nonlinear. Moreover, the theoretical analysis and the numerical algorithms of [52, 55] are completely based on this linearity property.

7.1.3 Contributions

First, we adapt the image convex model of [20, 30, 33, 91, 152] to *manifolds*. Then we show how to implement the method when the manifold is represented by a triangular *mesh*. Finally, we explain how our multi-region segmentation method could be incorporated into potential computer vision applications such as *3D reconstruction*.

7.2 Multi-Region Segmentation Model

In this section, we describe the convex image segmentation model we propose. To make this model comprehensible and intuitive, let us first remind of the region-based active contour model of Chan and Vese [31]. Here we show that the energy functional of Chan and Vese, which is the piecewise constant case of the Mumford-Shah model ([106]), can be recast as a convex functional in order to find the global minimizer of the original energy functional.

7.2.1 Convex Two-Phases Model

The Chan-Vese model [31], which is formulated in the level set framework, partition a given image into two subregions. For a given image I, the idea is to find a subset Σ of a bounded domain $\Omega \subset \mathbb{R}^N$, whose boundary $\partial\Sigma$ is represented by the zero level set of function $\phi : \Omega \rightarrow$

$\mathbb{R}^N$. This is done by minimizing the energy functional

$$\min_{\phi,\mu_1,\mu_2} \int_\Omega \left\{ H_\varepsilon(\phi)(I(\mathbf{w})-\mu_1)^2 + (1-H_\varepsilon(\phi))(I(\mathbf{w})-\mu_2)^2 + \lambda\,|\nabla H_\varepsilon(\phi)| \right\} d\omega\,, \qquad (7.1)$$

where $\lambda \in \mathbb{R}$, $\mu_1, \mu_2 \in \mathbb{R}$ and H_ε is a regularized Heaviside function, which models a characteristic function (see [31]).

Since the energy functional (7.1) is not convex, minimizing it by gradient descent methods can get stuck in local minima. By relaxing the characteristic function $H_\varepsilon(\phi)$ by an arbitrary function u bounded between 0 and 1, Chan *et al.* [33] showed that minimizing (7.1) can be rewritten as the following convex minimization problem:

$$\min_{0\le u\le 1} \left\{ \int_\Omega \left\{ u(\mathbf{w})(I(\mathbf{w})-\mu_1)^2 + (1-u(\mathbf{w}))(I(\mathbf{w})-\mu_2)^2 \right\} d\omega + \lambda \int_\Omega |\nabla u|\, d\omega \right\}, \qquad (7.2)$$

μ_1 and μ_2 being fixed, in $\mathbb{R}$. As proven in [20, 33], if $u(x)$ is a minimizer of (7.2), then for a.e. $\mu \in [0,1]$, the set $\Sigma(\mu) = \{x \in \Omega, u(x) > \mu\}$ is a minimizer of the Mumford-Shah functional [106], implying that *the solution to (7.1) can be obtained by thresholding u at any arbitrary threshold between* 0 *and* 1.

7.2.2 Extension to Multi-Region Segmentation

Recently, several authors [30, 91, 152] have extended the convex formulation (7.2) to multi-region segmentation:

$$\min_{\mathbf{u}\in K} \left\{ \int_\Omega < \mathbf{u}(\mathbf{w}), \mathbf{s}(\mathbf{w}) > + \lambda\ |\nabla \mathbf{u}(\mathbf{w})|\, d\omega \right\}, \qquad (7.3)$$

where K is the set of function $\mathbf{u} : \Omega \to \mathbb{R}^m$ such that for all $\mathbf{w} \in \Omega$ and $p \in [1..m]$, $\mathbf{u}_p(\mathbf{w}) \ge 0$ and $\sum_{p=1}^m \mathbf{u}_p(\mathbf{w}) = 1$. $|\nabla \mathbf{u}(\mathbf{w})|$ corresponds to $\sqrt{\sum_p |\nabla \mathbf{u}_p(\mathbf{w})|^2}$, where $|.|$ denotes the L^2 norm. m denotes the number of labels and $\mathbf{s}(\mathbf{w})$ is an m-dimensional vector; $s_p(\mathbf{w})$ indicates the affinity of the data at point $\mathbf{w}$ with class p. The convex domain naturally allow direct competition between the labeling.

7.3 Multi-Region Segmentation on Manifolds

In this section we extend the multi-region convex model (7.3) on a manifold, and we show how to optimize the associated energy for a manifold represented by a mesh. To our best

knowledge, these convex formulations (7.2,7.3) have been defined only on open subsets of $\mathbb{R}^N$ which correspond to image domains, as described in the previous section.

Let $\mathcal{S}$ be a Riemannian manifold. Typically, $\mathcal{S}$ could be a smooth 2D surface of $\mathbb{R}^3$. Energy (7.3) is adapted as follows:

$$\min_{\mathbf{u}\in K}\left\{ \int_{\mathcal{S}} < \mathbf{u}(\mathbf{x}), \mathbf{s}(\mathbf{x}) > + \lambda |\nabla_{\mathcal{S}} \mathbf{u}(\mathbf{x})| \, d\sigma \right\}, \tag{7.4}$$

where now the functions $\mathbf{u}$ are defined on $\mathcal{S}$ instead of Ω, $|.|$ is the Riemannian norm, $\nabla_{\mathcal{S}}$ is the intrinsic gradient on $\mathcal{S}$ and $d\sigma$ is the manifold's element measure (surface's area measure for 2D manifolds).

Now, let us consider a manifold represented by a mesh. The following results apply to manifolds with any topology. Let $\mathbf{X}$ be a (piecewise linear) polyhedron representation of the surface $\mathcal{S}$, defined by a set of vertices $\mathbf{x}_k : \mathbf{X} = \{\mathbf{x}_k\}$ and let l be the cardinality of $\mathbf{X}$ (the number of vertices). As in the finite elements literature, we define $\phi_k : \mathcal{S} \to \mathbb{R}$ as the piecewise affine, interpolating basis function such that $\phi_k(\mathbf{x}_k) = 1$ and $\phi_k(\mathbf{x}_i) = 0$ if $i \neq k$. The vector valued field $\mathbf{U} = \{\mathbf{u}_k\}$ is defined on all vertices $\mathbf{x}$ of the polyhedron $\mathbf{X}$. $\mathbf{U}$ can be naturally extended on $\mathcal{S}$ by a piecewise affine vector valued field on $\mathcal{S}$. We denote this extension $\mathbf{u}(\mathbf{x}) = \sum_k \mathbf{u}_k \phi_k(\mathbf{x})$. To make it easier to read, we assume that the manifold is a 2D surface of $\mathbb{R}^3$. However, the following method applies to any dimension. Let $\mathcal{S}_j$ be the j^{th} triangle of the mesh. The multi-region segmentation energy can then be rewritten as

$$\sum_j \sum_k \left\langle \mathbf{u}_k, \int_{\mathcal{S}_j} \phi_k(\mathbf{x}) \mathbf{s}(\mathbf{x}) d\sigma \right\rangle + \lambda \int_{\mathcal{S}_j} |\nabla_{\mathcal{S}} \mathbf{u}(\mathbf{x})| d\sigma, \tag{7.5}$$

where $\mathbf{u}$ is constrained to be in K. The first term of (7.5) is explicitly written with respect to $\mathbf{U}$. In order to make the total variation term explicit with respect to $\mathbf{U}$, we first consider a local parametrization (α, β) on the manifold. Following [54, 79], we rewrite the second term of Equation (7.5) using fundamental forms:

$$\nabla_{\mathcal{S}} \mathbf{u} = \begin{bmatrix} \frac{\partial \mathbf{x}}{\partial \alpha} & \frac{\partial \mathbf{x}}{\partial \beta} \end{bmatrix} \begin{bmatrix} E & F \\ F & G \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u}_\alpha \\ \mathbf{u}_\beta \end{bmatrix} \text{ , and then } |\nabla_{\mathcal{S}} \mathbf{u}| = \sqrt{\begin{bmatrix} \mathbf{u}_\alpha & \mathbf{u}_\beta \end{bmatrix} \begin{bmatrix} E & F \\ F & G \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u}_\alpha \\ \mathbf{u}_\beta \end{bmatrix}} \text{ ,}$$

where $\mathsf{E} = \frac{\partial \mathbf{x}}{\partial \alpha} \cdot \frac{\partial \mathbf{x}}{\partial \alpha}$, $\mathsf{F} = \frac{\partial \mathbf{x}}{\partial \alpha} \cdot \frac{\partial \mathbf{x}}{\partial \beta}$ and $\mathsf{G} = \frac{\partial \mathbf{x}}{\partial \beta} \cdot \frac{\partial \mathbf{x}}{\partial \beta}$ are coefficients of the first fundamental form (see [54, 79]). $\mathbf{u}_\alpha$ and $\mathbf{u}_\beta$ are partial derivatives of $\mathbf{u}$ with respects to α and β respectively. Considering the mesh representation, we parameterize the triangle $\mathcal{S}_j$ by $\mathbf{x}(\alpha, \beta) =$

$\mathbf{x}_{j,1} + \alpha \overrightarrow{\mathbf{x}_{j,1}\mathbf{x}_{j,2}} + \beta \overrightarrow{\mathbf{x}_{j,1}\mathbf{x}_{j,3}}$ where $\mathbf{x}_{j,1}$, $\mathbf{x}_{j,2}$ and $\mathbf{x}_{j,3}$ are the three vertices associated with the triangle $\mathcal{S}_j$ and where $(\alpha,\beta) \in T = \{(\alpha,\beta) | \alpha \in [0,1] \text{ and } \beta \in [0, 1-\alpha]\}$. We then have

$$\int_{\mathcal{S}_j} |\nabla_{\mathcal{S}} \mathbf{u}(\mathbf{x})| d\sigma = \int_T \sqrt{\sum_p \mathbf{u}_{p\alpha}^2 \mathsf{G} - 2\mathbf{u}_{p\alpha} \cdot \mathbf{u}_{p\beta} \mathsf{F} + \mathbf{u}_{p\beta}^2 \mathsf{E}}\ d\alpha d\beta\ . \tag{7.6}$$

$\mathbf{u}_{p\alpha}$ and $\mathbf{u}_{p\beta}$ are partial derivatives of $\mathbf{u}_p$ with respects to α and β respectively. Here the reader will easily verify that E, F, G, $\mathbf{u}_{p\alpha}$ and $\mathbf{u}_{p\beta}$ are constant functions on $\mathcal{S}_j$ and that their respective values are equal to $\mathsf{E}_j = |\mathbf{x}_{j,2} - \mathbf{x}_{j,1}|^2$, $\mathsf{F}_j = < \mathbf{x}_{j,2} - \mathbf{x}_{j,1}, \mathbf{x}_{j,3} - \mathbf{x}_{j,1} >$, $\mathsf{G}_j = |\mathbf{x}_{j,3} - \mathbf{x}_{j,1}|^2$, $\mathbf{u}_{p\alpha}^j = \mathbf{u}_{j,2_p} - \mathbf{u}_{j,1_p}$ and $\mathbf{u}_{p\beta}^j = \mathbf{u}_{j,3_p} - \mathbf{u}_{j,1_p}$, where $\mathbf{u}_{j,1}$, $\mathbf{u}_{j,2}$ and $\mathbf{u}_{j,3}$ are the values of $\mathbf{u}$ at vertices $\mathbf{x}_{j,1}$, $\mathbf{x}_{j,2}$ and $\mathbf{x}_{j,3}$ respectively. Now the term inside the integral of (7.6) does not depend on α and β. The convex multi-region segmentation energy on the meshed manifold becomes:

$$E(\mathbf{U}) = \sum_j \sum_k \left\langle \mathbf{u}_k, \int_{\mathcal{S}_j} \phi_k(\mathbf{x}) s(\mathbf{x}) d\sigma \right\rangle + \frac{\lambda}{2} \sum_j \sqrt{\sum_p {\mathbf{u}_{p\alpha}^j}^2 \mathsf{G}_\mathsf{j} - 2\mathbf{u}_{p\alpha}^j \cdot \mathbf{u}_{p\beta}^j \mathsf{F}_\mathsf{j} + {\mathbf{u}_{p\beta}^j}^2 \mathsf{E}_\mathsf{j}}\ . \tag{7.7}$$

7.3.1 Optimization Method

When the (surface) manifold is represented by a mesh, the convex multi-region segmentation model then leads to optimizing the convex energy (7.7) with respect to $\mathbf{U} \in \mathbb{R}^{l \times m}$, with the convex constraint $\mathbf{U} \in K$; K being the set $\{\mathbf{U} \text{ s.t. } \forall k,\ \sum_p \mathbf{u}_{kp} = 1 \text{ and } \forall p,\ \mathbf{u}_{kp} \geq 0\}$. This convex constrained optimization problem on $\mathbb{R}^{l \times m}$ can be solved by the projected gradient method [16], which consists in generating the sequence U^t via

$$\mathbf{U}^{t+1} = Proj_K(\mathbf{U}^t - \tau \nabla E(\mathbf{U}^t))\ , \tag{7.8}$$

for a fixed time step $\tau > 0$, until $|\mathbf{U}^t - \mathbf{U}^{t-1}|_\infty \leq \delta$, a small constant. $Proj_K$ is the projection on the convex set K. In other words, we iteratively process gradient descent steps and projections of the $\mathbf{u}_k$ on the set K. These projections can be done via Michelot's algorithm [99]. From energy (7.7) we easily obtain

$$\frac{\partial E}{\partial \mathbf{u}_{kp}}(\mathbf{U}) = \sum_{j \in N(k)} \left[\int_{\mathcal{S}_j} \phi_k(\mathbf{x}) s(\mathbf{x}) d\sigma \right]_p - \frac{\lambda}{2}\ Q\ (\xi + \varepsilon)^{-\frac{1}{2}}, \tag{7.9}$$

where $Q = (\mathbf{u}_{j,2_p} - \mathbf{u}_{kp})(\mathsf{G}_j - \mathsf{F}_j) + (\mathbf{u}_{j,3_p} - \mathbf{u}_{kp})(\mathsf{E}_j - \mathsf{F}_j)$, ξ is the term in the squared root of (7.7), and $N(k)$ is the 1-ring neighborhood of vertex k. As in [33], we regularize the term ξ by

incorporating a small value ε inside the squared root to avoid instabilities when the gradient of $\mathbf{u}$ is 0.

Let us remind now that, as underlined by [56], the notion of gradient depends on the underlying scalar product. If we chose the pointwise scalar product $<\mathbf{U},\mathbf{V}>_{\mathsf{pw}}=\sum_k <\mathbf{u}_k,\mathbf{v}_k>$, then the components of $\nabla E(\mathbf{U}^t)$ directly coincide with $\frac{\partial E}{\partial \mathbf{u}_{kp}}(\mathbf{U})$. Nevertheless the associate pointwise metric is not efficient for minimizing energies of the form $\int_S f(\mathbf{u}(\mathbf{x}))\, d\sigma$ since the distance between two discrete fields $\mathbf{U}$ and $\mathbf{V}$ does not take into account the area of the triangle. On the other hand, the L^2 scalar product $<\mathbf{U},\mathbf{V}>_{L^2}=\int_S <\mathbf{u}(\mathbf{x}),\mathbf{v}(\mathbf{x})> d\sigma$ is much more relevant. Also in this case the gradient becomes

$$\nabla E(\mathbf{U}) = M^{-1}\frac{\partial E}{\partial \mathbf{U}}(\mathbf{U})\,, \tag{7.10}$$

where the matrix M is the mass matrix defined by $M_{ij} = Id_m \int_S \phi_i(x)\phi_j(x)\, d\sigma$. Moreover one classically approximates M by the diagonal mass lumping $\tilde{M}$, where $\tilde{M}_{ii}$ is the area of the Voronoi dual cell of $\mathbf{x}_i$ times the identity matrix Id_m, see e.g. [56].

7.3.2 Applications

In the previous sections the data term of the segmentation model s is assumed to be known (Equation 7.4). In the applications, this term also depends on some parameters that have to be optimized. The convex problem can be solved by alternating optimization of the parameters in a bi-convex way. For fixed parameters of s we update $\mathbf{u}$ and vice-versa. $\mathbf{u}$ is updated according to the method presented in Section 7.3.1. In practice we update the parameters of s every r update iterations of $\mathbf{u}$ (r is chosen arbitrary; we fix $\mathsf{r}=10$ in our experiments).

Piecewise Constant Data Segmentation

Let us consider the case where the data we want to segment are assumed to be piecewise constant. Here a natural expression for $s_p(\mathbf{x})$ is to use the squared error between the scalar or vector-valued data $C(\mathbf{x})$ at the point $\mathbf{x}$ and the value μ_p associated with the label p (μ_p having the same dimension as the data):

$$s_p(\mathbf{x}) = (C(\mathbf{x})-\mu_p)^T(C(\mathbf{x})-\mu_p)\,.$$

The optimization of the energy (7.4) with respect to μ_p gives:

$$\mu_p = \frac{\int_{\mathbb{S}} \mathbf{u}_p(\mathbf{x})\, C(\mathbf{x}) d\sigma}{\int_{\mathbb{S}} \mathbf{u}_p(\mathbf{x}) d\sigma} ,$$

which corresponds to the mean value of the data of the associated region. Note that the previous model can be easily extended to any probability density function D_p. For example, D_p can be a multivariate gaussian density function of mean μ_p and covariance Σ_p, and then we would have:

$$s_p(\mathbf{x}) = -\ln(D_p(\mathbf{x}, \mu_p, \Sigma_p)) \text{ , with}$$

$$D_p(\mathbf{x}, \mu_p, \Sigma_p) = \frac{1}{\sqrt[m]{2\pi|\Sigma|}} e^{-\frac{1}{2}(C(\mathbf{x})-\mu_p)^T \Sigma_p^{-1} (C(\mathbf{x})-\mu_p)} .$$

Segmentation in 3D Reconstruction Problems

Such segmentation framework can be incorporated in 3D Reconstruction applications. In such applications, it can be interesting to segment a particular region, or all parts of the surface sharing the same reflectance properties. In 3D reconstruction, most of the variational methods yield to minimizing an energy of the form

$$E(\mathbb{S}) = \sum_i \int_{\mathbb{S}} g(\mathbf{x}) \frac{\mathbf{x}_i \cdot \mathbf{n}}{\mathbf{x}_{i,z}^3} \nu_{\mathbb{S}}(\mathbf{x})\, d\sigma , \tag{7.11}$$

see for example [49, 78, 79]. Moreover, if we choose

$$g(\mathbf{x}) = \sum_{p=1}^{m} \mathbf{u}_p(\mathbf{x}) (I_i(\pi_i(\mathbf{x})) - \mu_p)^T (I_i(\pi_i(\mathbf{x})) - \mu_p) ,$$

where $\pi_i(\mathbf{x})$ is the projection of the surface point $\mathbf{x}$ into the i^{th} image and $I_i : w \mapsto I_i(w)$ is the function which associates to each pixel w, its color on the i^{th} image. We then get an extension of the stereoscopic segmentation method proposed by [150] to the case where the surface is composed of more than two regions of piecewise constant radiance. Also, contrary to our method, the segmentation approach proposed in [150] is subject to local minima. Finally, the optimization of the energy (7.4) with respect to μ_p gives:

$$\mu_p = \frac{\int_{\mathbb{S}} \mathbf{u}_p(\mathbf{x})\ \sum_i I_i(\pi_i(\mathbf{x})) \frac{\mathbf{x} \cdot \mathbf{n}}{\mathbf{x}_z^3} \nu_{\mathbb{S},i}(\mathbf{x}) d\sigma}{\int_{\mathbb{S}} \mathbf{u}_p(\mathbf{x}) \sum_i \frac{\mathbf{x} \cdot \mathbf{n}}{\mathbf{x}_z^3} \nu_{\mathbb{S},i}(\mathbf{x}) d\sigma} .$$

If we chose

$$g(\mathbf{x}) = \sum_{p=1}^{m} \mathbf{u}_p(\mathbf{x})(I_i(\pi_i(\mathbf{x})) - \rho_p \mathbf{N}(\mathbf{x}) \cdot \mathbf{L})^T (I_i(\pi_i(\mathbf{x})) - \rho_p \mathbf{N}(\mathbf{x}) \cdot \mathbf{L}),$$

where $\mathbf{N}(\mathbf{x})$ is the normal to the surface at the point $\mathbf{x}$ and $\mathbf{L}$ is the vector corresponding to the light source illuminating the scene, then we get an extension of the (Lambertian) multi-view shape from shading method proposed by [78] for surfaces with piecewise constant albedo. In the same way, contrary to our approach, the method proposed by [78] is limited to two regions segmentation and is strongly subject to local minima. The optimization of energy (7.4) with respect to the albedo gives:

$$\rho_p = \frac{\int_{\mathcal{S}} \mathbf{u}_p(\mathbf{x}) \ \sum_i I_i(\pi_i(\mathbf{x})) \mathbf{N}(\mathbf{x}) \cdot \mathbf{L} \dfrac{\mathbf{x} \cdot \mathbf{n}}{\mathbf{x}_z^3} \nu_{\mathcal{S},i}(\mathbf{x}) d\sigma}{\int_{\mathcal{S}} \mathbf{u}_p(\mathbf{x}) (\mathbf{N}(\mathbf{x}) \cdot \mathbf{L})^2 \sum_i \dfrac{\mathbf{x} \cdot \mathbf{n}}{\mathbf{x}_z^3} \nu_{\mathcal{S},i}(\mathbf{x}) d\sigma} \ .$$

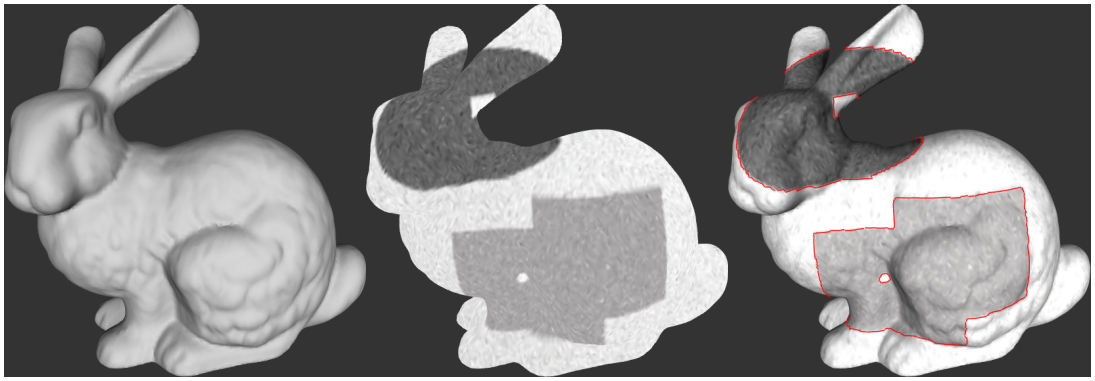
The theoretical and experimental study of these algorithms will be the topic a future work.

7.4 Experiments

In order to validate the proposed multi-region segmentation approach on meshes, we present different experiments on synthetic as well as realistic data. In practice as explained in previous section, the segmentation is solved by alternating between region parameters and the segmentation variable $\mathbf{U}$, with a known number of regions. The algorithm complexity is linearly dependent on the number of facets and the number of classes. Experiments have been run on a 2.66GHz linux machine and take about 20 seconds on a mesh of 200,000 facets and for a 4 regions segmentation. The values of λ have been manually chosen in each example but a value of 0.01 gives reasonable results in most cases.

7.4.1 The Two Region Case

Figures 7.2 and 7.3 show examples of our algorithm using a synthetic image mapped onto a mesh for the Stanford bunny model. Noise has been added to the image. Here, we show that our algorithm performs well on the given example and that the final solution is binary. Moreover it is robust to the initialization of the scalar function $\mathbf{U}$. Note that the retrieved solution that has been displayed is the auxiliary value $\mathbf{u}$, and not the segmented constant values



(a) (b) (c)

Figure 7.2: Segmentation result on the synthesized Stanford bunny surface. (a) Input shaded object. (b) Input mesh with synthetic texture mapping. (c) Input textured mesh (shaded visualization) and final contour (in red).

μ_1 and μ_2. Also because the energy functional is convex in $\mathbf{u}$ only and the values μ_1 and μ_2 are optimized during the evolution, they can be assigned to the region corresponding to either $\mathbf{u} = (1,0)$ or $\mathbf{u} = (0,1)$, this explains why the last initialization do not show the same values of $\mathbf{u}$ but an inverted one. In practice, although the total functional is not fully convex, we obtain the same results and really similar μ_1 and μ_2 for each example. We respectively obtain $(\mu_1 = 140.778\ ,\ \mu_2 = 231.003)$, $(\mu_1 = 140.746\ ,\ \mu_2 = 231.01)$, $(\mu_1 = 140.75\ , \mu_2 = 231.03\ )$ and $(\mu_1 = 230.992\ ,\ \mu_2 = 140.765)$ for the four different initializations. Note that in the last column, values of μ_1 and μ_2 are inverted and the solution $\mathbf{u}$ as well. In this example, geodesic active contours or level sets methods would tend to the closest local minima from the initialization as the texture is not clearly binary. Nevertheless, as the method is global here, segmenting a particular region should be done using additional cues.

Figure 7.4 present segmented surfaces from real-world textures that have been mapped onto a mesh, in the case of the two-phases segmentation. We show different examples from classical images used in segmentation. Note that the segmentation is done on the mesh using the method described in this chapter and not on an image. The experiments show three different non binary images and their segmentation into two different regions. As expected the results are binary even though the initial values of the segments are random values. The mean values of each region is estimated during the process as described before, and the parameters λ can be adjusted to add more smoothness to the segmentation. As shown by experiments, even

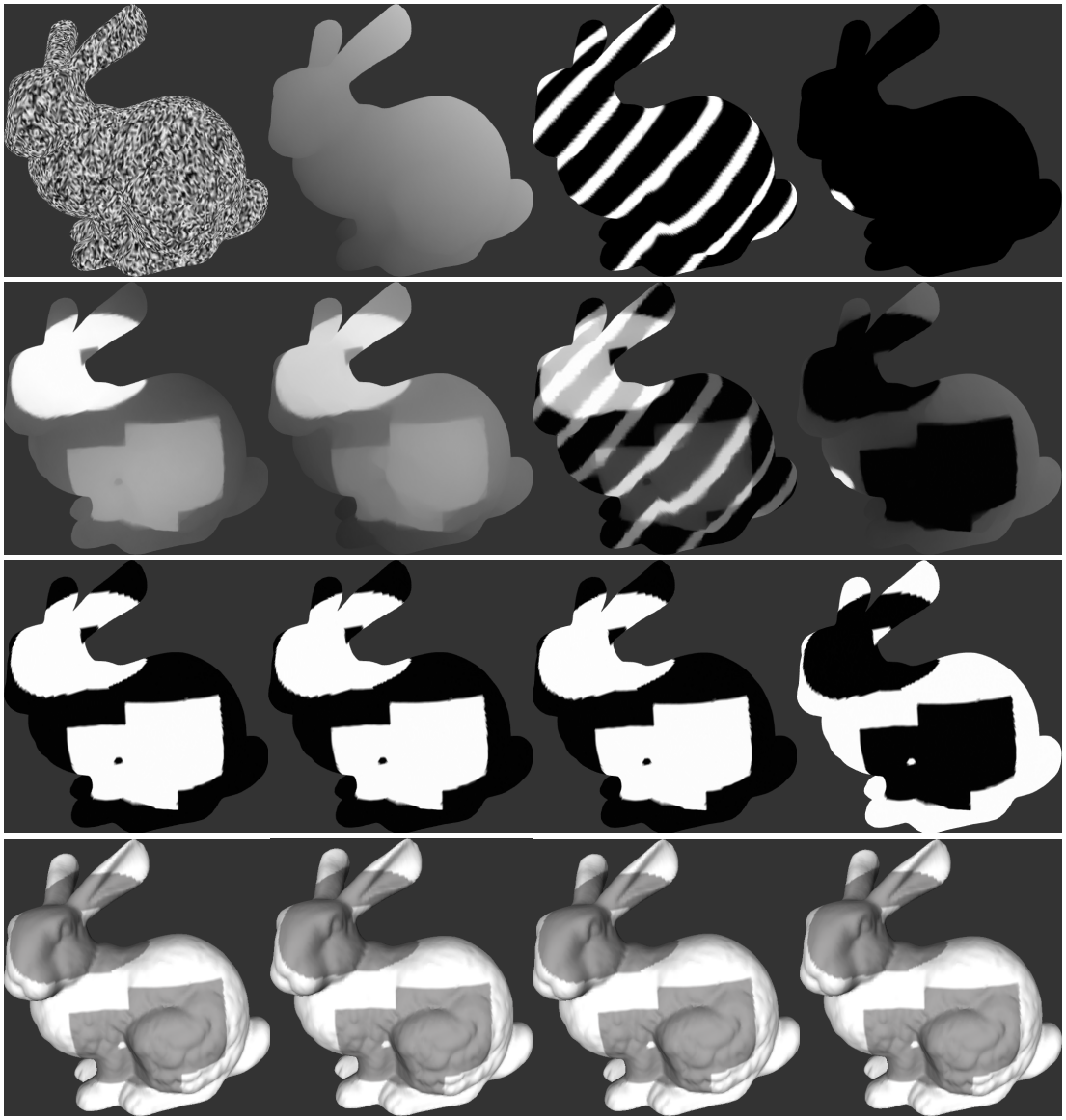


Figure 7.3: The evolution on the synthesized bunny surface. Different initialization of **U** (first row); Intermediate values of **U** (second row); The obtained solution **U** (third row); The obtained mean values (fourth row) with shading.

though the initialization is random and the parameters of each region are computed during the evolution, the algorithm still converges to the desired solution as a binary solution.

Figure 7.4: Segmentation results on meshes in the two-region segmentation on three different examples. From top to bottom: Input textured mesh; Mesh shape where the segmentation is performed and the initial random value of one component of $\mathbf{U}$; Recovered mean values of each region; Segmented object.

7.4.2 Dealing with Multiple Regions

Here we show the efficiency of the proposed method when dealing with multiple regions. Different examples are shown, first with synthetic textures on which noise has been added, and

then on meshes textured by real-world images like previous examples. Note that the number of regions is initially given and is not automatically estimated.

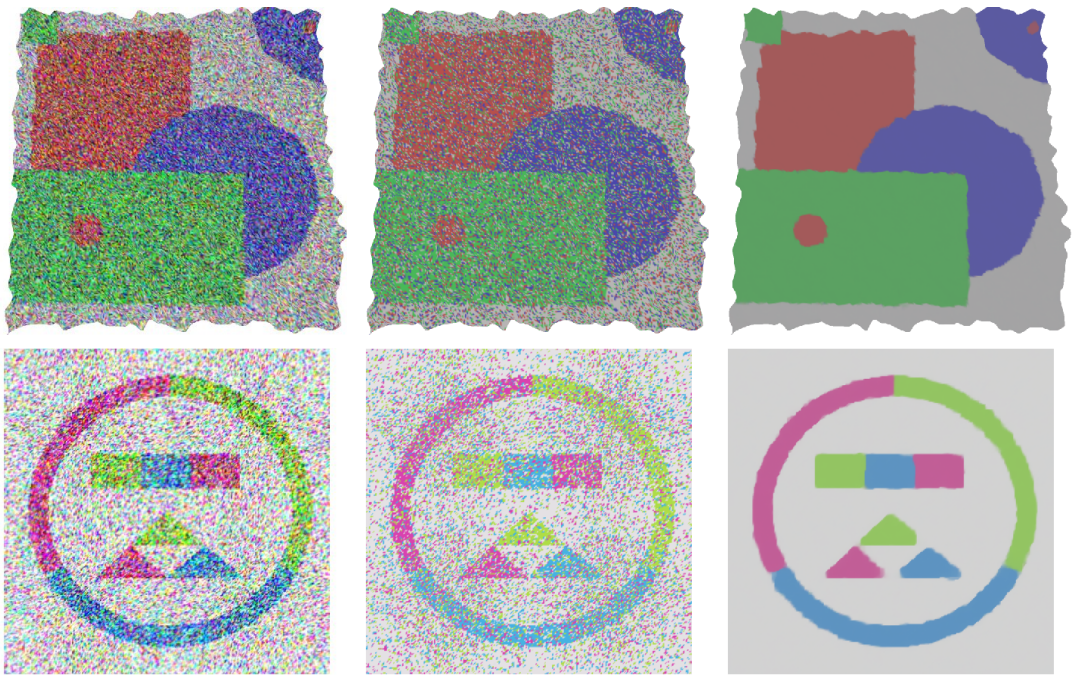


Figure 7.5: Segmentation results on meshes in the multi-region case on synthesized examples inspired from [91]. (a) Input textured mesh (same shape as previous Butterfly and Horses data). (b) Clustering using K-Means algorithm. (c) Recovered mean values of each region obtained by our approach.

In Figure 7.5, the experiment shows noisy texture on meshes, the segmentation result using K-Means, and the result of our TV-based algorithm on meshes. Because the K-Means algorithm does not take into account the spatial coherence of points, the result is noisy. On the other hand, the TV regularization allows coherence in the scene and the segmentation is close to the expected solution. In addition to be robust to initialization, our approach is robust to noise.

We then tested our multi-region segmentation approach on various data from real-world images [97], see Figure 7.6. Let us emphasize here that the initialization was random and the number of regions was initially given.

As an example, we applied our approach to segment mean curvature on a mesh using three different regions. Figure 7.7 shows that we are able to segment some concave and convex parts of the mesh.

Figure 7.6: Segmentation results on meshes in the multi-region case. Top row: Horse data set and its segmentation for three regions. Bottom row: Four regions labeling of the Butterfly data. (a) Input textured mesh; (b) Recovered mean values of each region obtained by our approach; (c) One of the segmented regions.

Finally, in Figure 7.8, we show the examples of a 3D mesh obtained by 3D reconstruction algorithms, as the one in [140, 153]. The last row shows the obtained color-based labeling (into three regions). Even though the texture is far from being binary, the segmentation is the expected one. For instance in the result, we nicely recover the skin, the pant and the shirt. Here again, initialization was random.

For comparison of the convex *image* multi-region segmentation model (7.3) with other methods, we refer to [91] which shows quantitative and qualitative comparisons with belief propagation, sequential belief propagation, graph cuts with alpha-expansion, graph cuts with alpha-beta swap and sequential tree reweighted belief propagation methods. The experiments show that the generated labeling is comparable to state-of-the-art discrete optimization methods.

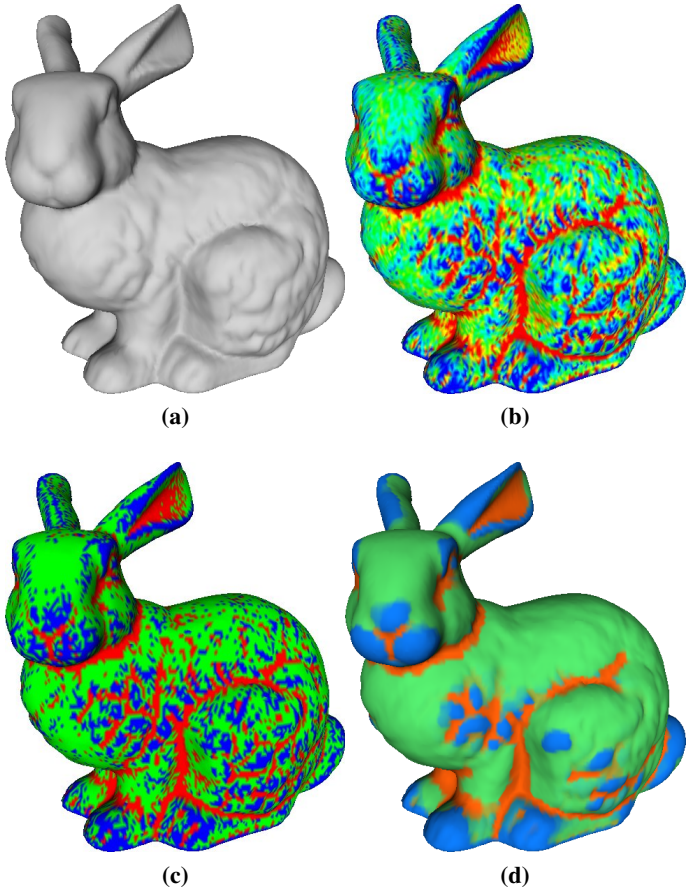


Figure 7.7: Segmentation results on mesh curvature in 3 regions. (a) Input mesh. (b) Mean Curvature visualization. (c) Simple thresholding of the mean curvature. (d) Segmentation result of the mean curvature into three regions with our approach.

7.5 Conclusions

In this chapter we propose a variational method for segmenting *data* on *manifolds* into regions of constant properties. The *convex* formulation makes the proposed model *robust to initialization*. Moreover, the total variation regularizer makes the method *robust to noise*. We show how to implement the method, in particular how to compute the gradient of the total variation term, when the surface has a discrete representation as triangular meshes. We have demonstrated the efficiency of our method by testing it on various synthetic and realistic data from computer vision applications.

Figure 7.8: Segmentation result on a colored mesh obtained by multiview stereo algorithm. Front view (top row) and back view (bottom row). Original input colored mesh and the associated 3D shape (left). Result of the segmentation into three regions obtained by our algorithm (middle). Recovered mean values displayed for each region (right).

Acknowledgements

We thank Andrei Zaharescu and Adrian Hilton for sharing the dancer data with the Perception group.

Chapter 8

Discussion and Future Work

Discussion

In this dissertation, we have presented prior-based variational segmentation models which incorporate low-level image features such as color, intensities and edges, together with higher-level visual knowledge such as occlusion information, smoothness constraint and shape priors to extract meaningful regions or objects from images, image sequences and manifolds. We have proposed using both local and global segmentation models of Chan and Vese with additional pose invariant shape prior constraints, which can be single known template, a result from a previous frame or a shape from shape manifolds.

First, we have presented a novel method for variational segmentation of image sequences containing nonrigid, moving objects. The proposed method is formulated as a variational problem, with one part of the functional corresponding to the Chan-Vese model and another part corresponding to the pose-invariant interaction term as a shape prior based on the previous contour. The optimal transformation as well as the shape deformation are determined by minimization of an energy functional using a gradient descent scheme. We have further coupled the proposed segmentation model with a registration process of two contours from the segmentation results of the previous frames in order to detect the occlusions, which is formulated as a variational contour matching problem. By using this occlusion information, the segmentation results can be further improved by reconstructing the occluded regions.

8. DISCUSSION AND FUTURE WORK

One of the main drawbacks of the variational segmentation models is that the energy functional is non-convex hence the initialization is critical in order to obtain global minima. Therefore, we have proposed a novel variational region-based active contour method for prior-based object segmentation in a global minimization framework. The method is based on convexifying the energy functional of the Chan-Vese method and adding a shape prior term as a constraint to segment an object whose global shape is given. We have shown and proven that the energy functional can be relaxed from optimization over characteristic functions to over arbitrary functions followed by a thresholding at an arbitrarily chosen level between 0 and 1.

As it is beneficial to have more sophisticated shape models in order to cope with the high variability of the object shape, we have presented a method that combines techniques from elastic shape manifold modeling, segmentation and optimization, to track the change of pose of a 3D object through tracking its contour. Both local and global models of the Chan-Vese with additional shape prior terms, have been proposed for the segmentation of the image sequences. It is worth to mention that although global models can be used in the view tracking application, local models that find local minima may be preferable.

The shape priors, used as additional terms in the segmentation models, need to be invariant with respect to a certain group of transformation. Since we use explicit manners to obtain the pose parameters, we have to determine an appropriate time step for each parameter to guarantee the stability which needs a careful tuning process. To overcome the common numerical problems associated with step size, we have proposed a new pose invariant shape prior model. This leads to a minimization problem which can be solved using continuous cut methods and to a derivation of the gradient expressions that is used as the basis for a stable and efficient gradient descent scheme for prior pose optimization.

Finally, we have proposed to a variational method for segmenting data on manifolds into several regions of constant properties. The convex formulation makes the proposed model robust to initialization and to noise. As many real applications, such as in Computer Graphics, use data which are represented by meshes to model objects, we have implemented the method to segment surfaces, which have discrete representations such as triangular meshes. We have then demonstrated the efficiency of the method by experiments on various synthetic and real data from computer vision applications.

Future work

It is worth to mention several directions of future work. First we would like to incorporate more prior knowledge into the segmentation models, such as motion as in [25] and texture information based on the Beltrami framework of Sochen *et al.* [136] as used in [75], to improve the robustness of the models. For segmentation of image sequences, it is reasonable to use dynamics and segmentation results from several previous frames instead of one previous frame as been proposed. Since it is necessary to handle the occlusions, we need to use better occlusion models and to use the spatial order information in the image model to dynamically impose prior only to occluded boundaries such as in [145]. We would also like to investigate more sophisticated shape models, especially non-rigid objects, which can better handle the variability of the shape of the objects such as in [34, 58, 59] or by using multi-reference shape priors as proposed in [62]. Regarding the drawbacks of optimization to obtain optimal transformation parameters in explicit manner, it would be interesting to investigate the use of implicit pose invariant such as in [41, 61]. So far we have used piecewise constant Mumford-Shah model which approximate the regions with constant values. Then it would be interesting to investigate to use piecewise smooth Mumford-Shah model which is more reasonable, instead of piecewise constant one, such as the work of Pock *et al.* [123] that proposed an algorithm for minimizing the piecewise smooth Mumford-Shah functional which is obtained by using functional lifting in higher dimension and convex relaxation. Another possible extension is to use a non-local extension of Total Variation such as in [19, 26, 71] which allows more flexibility in the regularization.

For convex multi-region segmentation problems, there are several directions of research. Firstly, since most of the proposed models so far cannot guarantee globally optimal solutions, therefore we would like to address this issue by investigating recent work in [7, 24] which are based on using a specific representation of the problem due to Lie *et al.* [94, 95]. As another important improvement regarding the numerical schemes for solving Total Variation based models, it would be interesting to further investigate the use of fast numerical schemes based on dual formulation such as in [8]. Finally, we consider the integration of shape priors to segment multi-objects of interest, for instance the work by Vu and Manjunath [149] in a discrete setting, into the convex multi-region models.

References

[1] E. AGANJ, J. P. PONS, F. SEGONNE, AND R. KERIVEN. **Spatio-temporal shape from silhouette using four-dimensional Delaunay meshing**. In *IEEE International Conference on Computer Vision*, Rio de Janeiro, Brazil, Oct 2007. 97

[2] L. AMBROSIO AND V.M. TORTORELLI. **Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence**. *Comm. on Pure and Applied Math.*, **XLIII**:999–1036, 1990. 7

[3] P. R. ANDRESEN AND M. NIELSEN. **Non-rigid Registration by Geometry-Constrained Diffusion**. In C. TAYLOR AND ET AL, editors, *MICCAI'99*, LNCS 1679, pages 533–543. Springer Verlag, 1999. 19, 51

[4] BEN APPLETON AND HUGUES TALBOT. **Globally Minimal Surfaces by Continuous Maximal Flows**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**(1), January 2006. 60, 96

[5] MARCO ATTENE, SAGI KATZ, MICHELA MORTARA, GIUSEPPE PATANÈ, MICHELA SPAGNUOLO, AND AYELLET TAL. **Mesh segmentation - A comparative study**. *Shape Modeling International*, 2006. 97

[6] GILLES AUBERT, MICHEL BARLAUD, OLIVIER D. FAUGERAS, AND STÉPHANIE JEHAN-BESSON. **Image Segmentation Using Active Contours: Calculus of Variations or Shape Gradients?** *SIAM Journal of Applied Mathematics*, **63**(6):2128–2154, 2003. 5

[7] E. BAE AND X. C. TAI. **Graph Cut Optimization for the Piecewise Constant Level Set Method Applied to Multiphase Image Segmentation**. In *Scale-Space and Variational Methods*, pages 1–13, 2009. 115

[8] EGIL BAE, JING YUAN, AND XUE-CHENG TAI. **Global Minimization for Continuous Multiphase Partitioning Problems using a Dual Approach**. Technical Report 09-75, Department of Mathematics, UCLA, September 2009. 115

[9] C. BALLESTER, V. CASELLES, AND J. VERDERA. **A variational model for disocclusion.** In *Proceeding ICIP (3)*, pages 677–680, 2003. 52

[10] MIKHAIL BELKIN AND PARTHA NIYOGI. **Laplacian Eigenmaps for Dimensionality Reduction and Data Representation**. *Neural Computation*, **15**(6):1373–1396, 2003. 68

[11] B. BERKELS. **An Unconstrained Multiphase Thresholding Approach for Image Segmentation**. In *Scale-Space and Variational Methods*, pages 26–37, 2009. 12

[12] M. BERTALMIO, G. SAPIRO, LI-TIEN CHENG, AND S. OSHER. **Variational problems and PDEs on implicit surfaces**. *Variational and Level Set Methods in Computer Vision*, 2001. 98

[13] A. BLAKE AND M. ISARD. *Active Contours*. Springer-Verlag, 1998. 4

[14] ANDREW BLAKE AND ANDREW. ZISSERMAN. *Visual reconstruction*. MIT Press, Cambridge, Mass. :, 1987. 7

[15] BLAISE BOURDIN AND ANTONIN CHAMBOLLE. **Implementation of an adaptive finite-element approximation of the Mumford–Shah functional**. *Numerische Mathematik*, **85**(4):609–646, June 2000. 7

[16] S. BOYD AND L. VANDENBERGHE. *Convex Optimization*. Cambridge University Press, Cambridge, U.K., 2003. 101

[17] Y. BOYKOV AND V. KOLMOGOROV. **An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **26**(9):1124–1137, September 2004. 3

[18] Y.Y. BOYKOV AND M.-P. JOLLY. **Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images**. *IEEE International Conference on Computer Vision*, 2001. 96

[19] A. BRAIDES AND G. DAL MASO. **Non-local approximation of the Mumford-Shah functional**. **5**, pages 293–322, 1997. 115

[20] X. BRESSON, S. ESEDOGLU, P. VANDERGHEYNST, J. P. THIRAN, AND S. J. OSHER. **Fast Global Minimization of the Active Contour/Snake Model**. *Journal of Mathematical Imaging and Vision*, **28**(2):151–167, June 2007. 17, 60, 85, 96, 98, 99

[21] X. BRESSON, P. VANDERGHEYNST, AND J-P. THIRAN. **Multiscale active contours**. *International Journal of Computer Vision*, **70**(3):197–211, 2006. 16

[22] X. BRESSON, P. VANDERGHEYNST, AND J-P. THIRAN. **A Variational Model for Object Segmentation Using Boundary Information and Shape Prior driven by the Mumford-Shah Functional**. *International Journal of Computer Vision*, **68**(2):145–162, 2006. 15, 16, 26, 44

[23] XAVIER BRESSON. *Image Segmentation with Variational Active Contours*. PhD thesis, ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE, 2005. 5

[24] ETHAN S. BROWN, TONY F. CHAN, AND XAVIER BRESSON. **Convex Formulation and Exact Global Solutions for Multi-phase Piecewise Constant Mumford-Shah Image Segmentation**. Technical Report 09-66, Department of Mathematics, UCLA, July 2009. 12, 115

[25] T. BROX, A. BRUHN, AND J. WEICKERT. **Variational motion segmentation with level sets**. In *ECCV 2006*, **3951** of *LNCS*, pages 471–483. Springer Verlag, 2006. 115

[26] ANTONI BUADES, BARTOMEU COLL, AND JEAN-MICHEL MORELT. **Nonlocal Image and Movie Denoising**. *International Journal of Computer Vision*, **76**(2):123–139, 2008. 115

[27] V. CASELLES, R. KIMMEL, AND G. SAPIRO. **Geodesic Active Contours**. *International Journal of Computer Vision*, **22**(1), 1997. 3, 6, 26, 44, 60, 96

[28] A. CHAMBOLLE. **An Algorithm for Total Variation Minimization and Applications**. *Journal of Mathematical Imaging and Vision*, **20**(1-2):89–97, January 2004. 22, 83, 85

[29] A. CHAMBOLLE. **Total Variation Minimization and a Class of Binary MRF Models**. In *EMMCVPR*, pages 136–152, 2005. 85, 91

[30] A. CHAMBOLLE, D. CREMERS, AND T. POCK. **A Convex Approach for computing minimal partitions**. Technical Report R.I.649, CMAP CNRS, November 2008. 12, 97, 98, 99

[31] T. CHAN AND L. VESE. **Active Contour Without Edges**. *IEEE Transactions on Image Processing*, **10**(2):266–277, 2001. 3, 6, 7, 10, 11, 18, 20, 23, 26, 27, 44, 45, 60, 74, 75, 76, 82, 96, 98, 99

[32] T. CHAN AND W. ZHU. **Level Set Based Prior Segmentation**. In *Proceeding CVPR 2005*, **2**, pages 1164–1170, 2005. 15, 26, 32, 33, 34, 44, 46, 47, 83

[33] TONY F. CHAN, SELIM ESEDOGLU, AND MILA NIKOLOVA. **Algorithms for Finding Global Minimizers of Image Segmentation and Denoising Models**. *Journal of Applied Mathematics*, 2006. 10, 11, 12, 15, 20, 60, 61, 62, 63, 64, 74, 75, 76, 82, 84, 85, 86, 97, 98, 99, 101

[34] G. CHARPIAT, O. D. FAUGERAS, AND R. KERIVEN. **Approximations of shape metrics and application to shape warping and empirical shape statistics**. *Foundations of Computational Mathematics*, **5**(1):1–58, 2005. 115

[35] Y. CHEN, H. D. TAGARE, S. THIRUVENKADAM, F. HUANG, D. WILSON, K. S. GOPINATH, R. W. BRIGGS, AND E. A. GEISER. **Using prior shapes in geometric active contours in a variational framework**. *International Journal of Computer Vision*, **50**(3):315–328, 2002. 15, 16, 21, 26, 44, 60, 63, 74, 75

[36] L. D. COHEN AND R. KIMMEL. **Global Minimum for Active Contour Models: A Minimal Path Approach**. *International Journal of Computer Vision*, **24**(1):57–78, August 1997. 10

[37] T. F. COOTES, C. J. TAYLOR, D. H. COOPER, AND J. GRAHAM. **Active Shape Models: Their Training and Application**. *Computer Vision and Image Understanding*, **61**(1):38–59, 1995. 16

[38] D. CREMERS. **Dynamical Statistical Shape Priors for Level Set-Based Tracking**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**(8):1262–1273, 2006. 26

[39] D. CREMERS AND G. FUNKA-LEA. **Dynamical Statistical Shape Priors for Level Set Based Sequence Segmentation**. In N. PARAGIOS AND ET AL., editors, *3rd Workshop on Variational and Level Set Methods in Computer Vision*, LNCS 3752, pages 210–221. Springer Verlag, 2005. 15, 44

[40] D. CREMERS, T. KOHLBERGER, AND C. SCHNÖRR. **Shape statistics in kernel space for variational image segmentation**. *Pattern Recognition*, **36**:1929–1943, 2003. 16

[41] D. CREMERS, S. J. OSHER, AND S. SOATTO. **Kernel Density Estimation and Intrinsic Alignment for Shape Priors in Level Set Segmentation**. *International Journal of Computer Vision*, **69**(3):335–351, 2006. 16, 17, 26, 60, 115

[42] D. CREMERS AND S. SOATTO. **A Pseudo-distance for Shape Priors in Level Set Segmentation**. In O. FAUGERAS AND N. PARAGIOS, editors, *2nd IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision*, 2003. 15, 26, 32, 33, 35, 36, 44, 46, 47, 83

[43] D. CREMERS, N. SOCHEN, AND C. SCHNÖRR. **Towards Recognition-Based Variational Segmentation Using Shape Priors and Dynamic Labeling**. In L.D. GRIFFIN AND M. LILLHOLM, editors, *Scale Space 2003*, LNCS 2695, pages 388–400. Springer Verlag, 2003. 44

[44] D. CREMERS, N. SOCHEN, AND C. SCHNÖRR. **A multiphase dynamic labeling model for variational recognition-driven image segmentation**. *International Journal of Computer Vision*, **66**(1):67–81, 2006. 26

[45] D. CREMERS, F. TISCHHÄUSER, J. WEICKERT, AND C. SCHNÖRR. **Diffusion snakes: introducing statistical shape knowledge into the mumford-shah functional**. *International Journal of Computer Vision*, **50**(3):295–313, 2002. 4, 15

[46] S. DAMBREVILLE, Y. RATHI, AND A. TANNENBAUM. **A Framework for Image Segmentation Using Shape Models and Kernel Space Shape Priors**. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, **30**(8):1385–1399, August 2008. 16

[47] B. C. DAVIS, P. T. FLETCHER, E. BULLITT, AND S. JOSHI. **Population Shape Regression From Random Design Data**. In *ICCV*, pages 1–7, 2007. 68

[48] AMAËL DELAUNOY, KETUT FUNDANA, EMMANUEL PRADOS, AND ANDERS HEYDEN. **Convex Multi-Region Segmentation on Manifolds**. In IEEE, editor, *Proceedings of the 12th International Conference on Computer Vision, Kyoto, Japan*, sept 2009. vi, 18, 22, 95

[49] AMAËL DELAUNOY, EMMANUEL PRADOS, PAU GARGALLO, JEAN-PHILIPPE PONS, AND PETER STURM. **Minimizing the Multi-view Stereo Reprojection Error for Triangular Surface Meshes**. In *British Machine and Vision Conference, Leeds, UK*, 2008. 97, 103

[50] M. C. DELFOUR AND J.-P. ZOLESIO. *Shapes and Geometries. Analysis, Differential Calculus, and Optimization*. Advances in Design and Control. SIAM, 2001. 5, 36

[51] H. DELINGETTE AND J. MONTAGNAT. **Shape and Topology Constraints on Parametric Active Contours**. *Computer Vision and Image Understanding*, **83**(2):140–171, August 2001. xvii, 74, 77

[52] A. DEMLOW AND G. DZIUK. **An Adaptive Finite Element Method for the Laplace-Beltrami Operator on Implicitly Defined Surfaces**. *SIAM Journal of Numerical Analysis*, **45**(1), 2007. 98

[53] A. DERVIEUX AND F. THOMASSET. **A Finite Element Method for the Simulation of Rayleigh-Taylor Instability**. **771**:145–158, 1980. 4

[54] MANFREDO P. DO CARMO. *Differential Geometry of Curves and Surfaces*. Prentice Hall, February 1976. 100

[55] G. DZIUK. **Finite Elements for the Beltrami operator on arbitrary surfaces**. In SPRINGER BERLIN / HEIDELBERG, editor, *Partial Differential Equations and Calculus of Variations*, pages 142–155. Lecture Notes in Mathematics, 2006. 98

[56] I. ECKSTEIN, J.-P. PONS, Y. TONG, C.-C. J. KUO, AND M. DESBRUN. **Generalized surface flows for mesh processing**. In *Eurographics Symposium on Geometry Processing*, 2007. 102

[57] D. W. EGGERT, L. STARK, AND K. W. BOWYER. **Aspect Graphs and Their Use in Object Recognition**. *Annals of Mathematics and Artificial Intelligence*, **13**(3-4):347–375, 1995. 69

[58] P. ETYNGIER, F. SEGONNE, AND R. KERIVEN. **Shape Priors using Manifold Learning Techniques**. In *International Conference on Computer Vision*, pages 1–8, 2007. 16, 68, 115

[59] PATRICK ETYNGIER, FLORENT SÉGONNE, AND RENAUD KERIVEN. **Active-Contour-Based Image Segmentation Using Machine Learning Techniques**. In NICHOLAS AYACHE, SÉBASTIEN OURSELIN, AND ANTHONY J. MAEDER, editors, *MICCAI (1)*, **4791** of *Lecture Notes in Computer Science*, pages 891–899. Springer, 2007. 16, 115

[60] WENDELL H. FLEMING AND RAYMOND RISHEL. **An integral formula for total gradient variation**. *Archiv der Mathematik*, **11**:218–222, 1960. 10, 14

[61] A. FOULONNEAU, P. CHARBONNIER, AND F. HEITZ. **Affine-Invariant Geometric Shape Priors for Region-Based Active Contours**. *IEEE Trans. Pattern Analysis and Machine Intelligence*, **28**(8):1352–1357, August 2006. 15, 17, 115

[62] A. FOULONNEAU, P. CHARBONNIER, AND F. HEITZ. **Multi-Reference Shape Priors for Active Contours**. *International Journal of Computer Vision*, **81**(1):68–81, January 2009. 115

[63] JEAN FRANCOIS AUJOL AND ANTONIN CHAMBOLLE. **Dual Norms and Image Decomposition Models**. *Int. J. Comput. Vis.*, **63**(1):85–104, 2005. 85

[64] K. FUNDANA, A. HEYDEN, C. GOSCH, AND C. SCHNORR. **Continuous graph cuts for prior-based object segmentation**. In *Proc. 19th International Conference on Pattern Recognition ICPR 2008*, pages 1–4, December 8–11, 2008. vi, 18, 20, 59, 83, 86

[65] K. FUNDANA, N. C. OVERGAARD, AND A. HEYDEN. **Variational Segmentation of Image Sequences Using Region-Based Active Contours and Deformable Shape Priors**. *International Journal of Computer Vision*, **80**(3):289–299, December 2008. vi, 18, 25

[66] K. FUNDANA, N.C. OVERGAARD, AND A. HEYDEN. **Deformable Shape Priors in Chan-Vese Segmentation of Image Sequences**. In *IEEE International Conference on Image Processing 2007*, pages 285–288, 2007. vii, 18

[67] K. FUNDANA, N.C. OVERGAARD, AND A. HEYDEN. **Variational Segmentation of Image Sequences using Deformable Shape Priors**. In *SCIA 2007*, LNCS 4522, pages 31–40. Springer-Verlag, 2007. vii, 18, 31

[68] KETUT FUNDANA, NIELS CHR. OVERGAARD, ANDERS HEYDEN, DAVID GUSTAVSSON, AND MADS NIELSEN. **Nonrigid Object Segmentation and Occlusion Detection in Image Sequences**. In ALPESH RANCHORDAS AND HELDER ARAÚJO, editors, *VISAPP (1)*, pages 211–218. INSTICC - Institute for Systems and Technologies of Information, Control and Communication, 2008. vi, 18, 19, 43

[69] STUART GEMAN AND DONALD GEMAN. **Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6):721–741, November 1984. 7

[70] C. GENTILE, O. CAMPS, AND M. SZNAIER. **Segmentation for Robust Tracking in the Presence of Severe Occlusion**. *IEEE Transactions on Image Processing*, **13**(2):166–178, 2004. 26, 44

[71] GUY GILBOA AND STANLEY OSHER. **Nonlocal Operators with Applications to Image Processing**. *Multiscale Model. Simul.*, **7**(3):1005–1028, 2008. 115

[72] TOM GOLDSTEIN, XAVIER BRESSON, AND STANLEY OSHER. **Geometric Applications of the Split Bregman Method: Segmentation and Surface Reconstruction**. Technical Report 09-06, Department of Mathematics, UCLA, February 2009. 17

[73] C. GOSCH, K. FUNDANA, A. HEYDEN, AND C. SCHNORR. **View Point Tracking of Rigid Objects Based on Shape Sub-manifolds**. In *ECCV*, pages III: 251–263, 2008. vi, 18, 20, 67

[74] D. GUSTAVSSON, K. FUNDANA, N. C. OVERGAARD, A. HEYDEN, AND M. NIELSEN. **Variational Segmentation and Contour Matching of Non-Rigid Moving Object**. In *ICCV Workshop on Dynamical Vision 2007*, LNCS. Springer Verlag, 2007. vii, 18, 19, 43

[75] N. HOUHOU, J. P. THIRAN, AND X. BRESSON. **Fast texture segmentation model based on the shape operator and active contour**. In *Computer Vision and Pattern Recognition*, pages 1–8, 2008. 115

[76] HIROSHI ISHIKAWA. **Exact Optimization for Markov Random Fields with Convex Priors**. *IEEE Transaction Pattern Analysis and Machine Intelligence*, **25**, 2003. 97

[77] STÉPHANIE JEHAN-BESSON, MICHEL BARLAUD, AND GILLES AUBERT. **DREAM2S: Deformable Regions Driven by an Eulerian Accurate Minimization Method for Image and Video Segmentation**. *International Journal of Computer Vision*, **53**(1):45–70, 2003. 6

[78] H. JIN, D. CREMERS, D. WANG, E. PRADOS, A. YEZZI, AND S. SOATTO. **3-D Reconstruction of Shaded Objects from Multiple Images Under Unknown Illumination**. *International Journal of Computer Vision*, **76**(3), March 2008. 96, 97, 103, 104

[79] H. JIN, A. YEZZI, Y.-H. TSAI, L.-T. CHENG, AND S. SOATTO. **Estimation of 3D Surface Shape and Smooth Radiance from 2D Images: A Level Set Approach**. *J. Sci. Comput.*, **19**(1-3), 2003. 100, 103

[80] S. H. JOSHI, E. KLASSEN, A. SRIVASTAVA, AND I. H. JERMYN. **An Efficient Representation for Computing Geodesics Between n-Dimensional Elastic Shapes**. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2007. 69

[81] S. H. JOSHI, E. KLASSEN, A. SRIVASTAVA, AND I. H. JERMYN. **Removing Shape-Preserving Transformations in Square-Root Elastic (SRE) Framework for Shape Analysis of Curves**. In *EMMCVPR*, pages 387–398, 2007. 70

[82] S. H. JOSHI AND A. SRIVASTAVA. **Intrinsic Bayesian Active Contours for Extraction of Object Boundaries in Images**. *International Journal of Computer Vision*, **81**(3):331–355, March 2009. 68, 76

[83] M. KASS, A. WITKIN, AND D. TERZOPOULOS. **Snakes: Active Contour Models**. *International Journal of Computer Vision*, pages 321–331, 1988. 3, 5, 26, 60, 96

[84] R. KIMMEL AND A. M. BRUCKSTEIN. **Regularized Laplacian Zero Crossings as Optimal Edge Integrators**. *International Journal of Computer Vision*, **53**(3):225–243, July 2003. 6

[85] ERIC KLASSEN, ANUJ SRIVASTAVA, WASHINGTON MIO, AND SHANTANU H. JOSHI. **Analysis of Planar Shapes Using Geodesic Paths on Shape Spaces**. *IEEE Transaction Pattern Analysis and Machine Intelligence*, **26**(3):372–383, 2003. 67, 68

[86] P. KOHLI AND P. H. S. TORR. **Dynamic Graph Cuts for Efficient Inference in Markov Random Fields**. *IEEE Transaction Pattern Analysis and Machine Intelligence*, December 2007. 96

[87] KOLMOGOROV AND ZABIH. **What Energy Functions Can Be Minimized via Graph Cuts**. *IEEE Transaction Pattern Analysis and Machine Intelligence*, **26**, 2004. 3, 96

[88] J. KONRAD AND M. RISTIVOJEVIC. **Video segmentation and occlusion detection over multiple frames**. In B. VASUDEV, T. R. HSING, A. G. TESCHER, AND T. EBRAHIMI, editors, *Image and Video Communications and Processing 2003*, SPIE 5022, pages 377–388. SPIE, 2003. 44

[89] MATTHIAS KRUEGER, PATRICE DELMAS, AND GEORGY GIMEL'FARB. **Active Contour Based Segmentation of 3D Surfaces**. In *European Conference on Computer Vision*, 2008. 97

[90] C. S. LEE AND A. M. ELGAMMAL. **Modeling View and Posture Manifolds for Tracking**. In *ICCV*, pages 1–8, 2007. 68

[91] J. LELLMANN, J. KAPPES, J. YUAN, F. BECKER, AND C. SCHNOERR. **Convex Multi-Class Image Labeling by Simplex-Constrained Total Variation**. Technical report, University of Heidelberg, oct 2008. xix, 97, 98, 99, 108, 109

[92] JAN LELLMANN, JÖRG H. KAPPES, JING YUAN, FLORIAN BECKER, AND CHRISTOPH SCHNÖRR. **Convex Multi-class Image Labeling by Simplex-Constrained Total Variation**. In XUE-CHENG TAI, KNUT MØRKEN, MARIUS LYSAKER, AND KNUT-ANDREAS LIE, editors, *SSVM*, **5567** of *Lecture Notes in Computer Science*, pages 150–162. Springer, 2009. 12, 13

[93] M.E. LEVENTON, W.E.L. GRIMSON, AND O. FAUGERAS. **Statistical Shape Influence in Geodesic Active Contours**. In *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pages 316–323, 2000. 16, 26, 44, 60, 83

[94] J. LIE, M. LYSAKER, AND X. C. TAI. **Piecewise Constant Level Set Methods and Image Segmentation**. In *Scale-Space and Variational Methods*, pages 573–584, 2005. 115

[95] JOHAN LIE, MARIUS LYSAKER, AND XUE CHENG TAI. **A Variant Of The Level Set Method And Applications to Image Segmentation**. *Mathematics of Computation*, **75**(255):1155–1174, 2006. 115

[96] B. LUCAS AND T. KANADE. **An iterative image registration technique with an application to stereo vision**. In *Proc. DARPA Image Understanding Workshop*, pages 121–130, 1981. 36

[97] D. MARTIN, C. FOWLKES, D. TAL, AND J. MALIK. **A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics**. In *International Conference on Computer Vision*, **2**, pages 416–423, July 2001. 64, 108

[98] LEVENTON M.E., FAUGERAS O., GRIMSON W.E.L., AND WELLS W.M. III. **Level set based segmentation with intensity and curvature priors**. pages 4–11, 2000. 14

[99] C MICHELOT. **A finite algorithm for finding the projection of a point onto the Canonical simplex of R^n**. *Journal of Optimization Theory and Applications*, **50**(1):195–200, 1986. 101

[100] PETER W. MICHOR, DAVID MUMFORD, JAYANT SHAH, AND LAURENT YOUNES. **A Metric on Shape Space with Explicit Geodesics**, May 05 2007. Comment: 31 pages, some typos corrected. 69

[101] W. MIO, A. SRIVASTAVA, AND S. H. JOSHI. **On Shape of Plane Elastic Curves**. *International Journal of Computer Vision*, **73**(3):307–324, July 2007. 20, 67, 68, 69

[102] W. MIO, A. SRIVASTAVA, AND X. W. LIU. **Contour Inferences for Image Understanding**. *International Journal of Computer Vision*, **69**(1):137–144, August 2006. 67

[103] M. MOELICH AND T. CHAN. **Tracking Objects with the Chan-Vese Algorithm**. Technical Report 03-14, Department of Mathematics, UCLA, March 2003. 16, 26, 37

[104] B. MORY, R. ARDON, AND J. P. THIRAN. **Variational Segmentation using Fuzzy Region Competition and Local Non-Parametric Probability Density Functions**. In *International Conference on Computer Vision*, pages 1–8, 2007. 11, 12

[105] D. MUMFORD AND J. SHAH. **Boundary Detection by Minimizing Functionals**. In *Proceeding of International Conference on Computer Vision and Pattern Recognition*, pages 22–26, 1985. 6, 7

[106] D. MUMFORD AND J. SHAH. **Optimal Approximations by Piecewise Smooth Functions and Variational Problems**. *Comm. on Pure and Applied Math.*, **XLII**(5):577–685, 1988. 3, 6, 7, 10, 27, 60, 61, 98, 99

[107] D. MUMFORD AND J. SHAH. **Finite Difference Approximation of the Mumford-Shah Functional**. *Comm. on Pure and Applied Math.*, (51):197–228, 1998. 7

[108] CARL OLSSON, MARTIN BYRÖD, NIELS CHR. OVERGAARD, AND FREDRIK KAL. **Extending Continuous Cuts: Anisotropic Metrics and Expansion Moves**. In *International Conference on Computer Vision*, 2009. 12

[109] N. U. OPTIMIZATION TECHNOLOGY CENTER. **The NEOS Guide**. http://www-fp.mcs.anl.gov/OTC/Guide/. 73

[110] S. OSHER AND R. FEDKIW. *Level Set Methods and Dynamic Implicit Surfaces*. Springer-Verlag, New York, 2003. 4, 28, 46, 76, 82

[111] S. OSHER AND JAMES A. SETHIAN. **Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations**. *Journal of Computational Physics*, **79**:12–49, 1988. 4, 6, 82

[112] N. C. OVERGAARD, K. FUNDANA, AND A. HEYDEN. **Pose Invariant Shape Prior Segmentation Using Continuous Cuts and Gradient Descent on Lie Groups**. In *Proceeding of International Conference on Scale-Space and Variational Methods*, pages 684–695, 2009. vi, 18, 21, 81

[113] D. L. PAGE, A. F. KOSCHAN, AND M. A. ABIDI. **Perception-based 3D Triangle Mesh Segmentation Using Fast Marching Watersheds**. *Computer Vision and Pattern Recognition*, 2003. 98

[114] N. PARAGIOS AND R. DERICHE. **Geodesic Active Contours and Level Set Methods for the Detection and Tracking of Moving Objects**. *IEEE Trans. PAMI*, **22**(3):266–280, 2000. 26

[115] N. PARAGIOS AND R. DERICHE. **Geodesic active regions: a new paradigm to deal with frame partition problems in computer vision**. *Journal of Visual Communication and Image Representation*, **13**(1-2):249–268, 2002. 3, 6

[116] N. PARAGIOS AND R. DERICHE. **Geodesic Active Regions and Level Set Methods for motion estimation and Tracking**. *Computer Vision and Image Understanding*, **97**:259–282, 2005. 26, 60, 96

[117] N. PARAGIOS, M. ROUSSON, AND VISVANATHAN RAMESH. **Matching Distance Functions: A Shape-to-Area Variational Approach for Global-to-Local Registration**. In A. HEYDEN AND ET AL, editors, *ECCV 2002*, LNCS 2351, pages 775–789. Springer-Verlag Berlin Heidelberg, 2003. 35, 36

[118] XAVIER PENNEC. **Probabilities and statistics on Riemannian manifolds: Basic tools for geometric measurements**. In A. ENIS ÇETIN, LALE AKARUN, AYSIN ERTÜZÜN, METIN N. GURCAN, AND YASEMIN YARDIMCI, editors, *NSIP*, pages 194–198. Bogaziçi University Printhouse, 1999. 72

[119] G. PEYRE AND L. COHEN. **Surface segmentation using geodesic centroidal tesselation**. *International Symposium on 3D Data Processing, Visualization and Transmission*, Sept. 2004. 97

[120] T. POCK, A. CHAMBOLLE, D. CREMERS, AND H. BISCHOF. **A convex relaxation approach for computing minimal partitions**. In *Internation Conference on Computer Vision and Pattern Recognition*, pages 810–817, 2009. 12

[121] T. POCK, T. SCHOENEMANN, G. GRABER, H. BISCHOF, AND D. CREMERS. **A Convex Formulation of Continuous Multi-label Problems**. In *European Conference on Computer Vision*, pages III: 792–805, 2008. 12, 97

[122] THOMAS POCK. *Fast Total Variation for Computer Vision*. Phd thesis, Graz University of Technology, 2008. 15

[123] THOMAS POCK, DANIEL CREMERS, HORST BISCHOF, AND ANTONIN CHAMBOLLE. **An Algorithm for Minimizing the Mumford-Shah Functional**. In *IEEE International Conference on Computer Vision (ICCV)*, 2009. 115

[124] J.-P. PONS AND J.-D. BOISSONNAT. **Delaunay deformable models: Topology-adaptive meshes based on the restricted Delaunay triangulation**. In *IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, USA, Jun 2007. 97

[125] RENFREY B. POTTS. **Some Generalized Order-disorder Transformations**. In *Proceeding of the Cambridge Philosophical Society*, **48**, pages 106–109, 1952. 12

[126] TAMMY RIKLIN-RAVIV, NAHUM KIRYATI, AND NIR SOCHEN. **Prior-based Segmentation and Shape Registration in the Presence of Perspective Distortion**. *International Journal of Computer Vision*, **72**(3):309–328, 2007. 16, 34, 60

[127] TAMMY RIKLIN-RAVIV, NAHUM KIRYATI, AND NIR A. SOCHEN. **Unlevel-Sets: Geometry and Prior-Based Segmentation**. In TOMÁS PAJDLA AND JIRI MATAS, editors, *ECCV (4)*, **3024** of *Lecture Notes in Computer Science*, pages 50–61. Springer, 2004. 21, 74, 83

[128] M. ROUSSON AND N. PARAGIOS. **Shape Priors for Level Set Representations**. In A. HEYDEN AND ET AL, editors, *ECCV 2002*, LNCS 2351, pages 78–92. Springer Verlag, 2002. 15, 16, 26, 35, 36, 44, 83

[129] M. ROUSSON AND N. PARAGIOS. **Prior Knowledge, Level Set Representations and Visual Grouping**. *International Journal of Computer Vision*, **76**(3):231–243, March 2008. 15, 26, 60

[130] MIKAËL ROUSSON AND DANIEL CREMERS. **Efficient Kernel Density Estimation of Shape and Intensity Priors for Level Set Segmentation**. In JAMES S. DUNCAN AND GUIDO GERIG, editors, *MICCAI (2)*, **3750** of *Lecture Notes in Computer Science*, pages 757–764. Springer, 2005. 15, 26

[131] S. T. ROWEIS AND L. K. SAUL. **Nonlinear Dimensionality Reduction by Locally Linear Embedding**. *Science*, **290**(5500):2323–2326, December 2000. 68

[132] L. I. RUDIN, S. J. OSHER, AND E. FATEMI. **Nonlinear total variation based noise removal algorithms**. *Physica D: Nonlinear Phenomena*, **60**:259–268, 1992. 10, 14, 85

[133] T. B. SEBASTIAN, P. N. KLEIN, AND B. B. KIMIA. **On Aligning Curves**. *IEEE Transaction Pattern Analysis and Machine Intelligence*, **25**:116–124, 2003. 69

[134] J.A. SETHIAN. *Level Set Methods and Fast Marching Methods Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. Cambridge University Press, 1999. 82

[135] PHILIP SHILANE, PATRICK MIN, MICHAEL M. KAZHDAN, AND THOMAS A. FUNKHOUSER. **The Princeton Shape Benchmark**. In *Shape Modeling International*, pages 167–178. IEEE Computer Society, 2004. xvii, 70

[136] N. A. SOCHEN, R. KIMMEL, AND R. MALLADI. **A General Framework for Low-Level Vision**. *IEEE Trans. Image Processing*, **7**(3):310–318, March 1998. 115

[137] J. E. SOLEM. **Geodesic Curves for Analysis of Continuous Implicit Shapes**. In *ICPR*, pages I: 43–46, 2006. 76

[138] J. E. SOLEM AND N. CHR. OVERGAARD. **A geometric formulation of gradient descent for variational problems with moving surfaces**. In R. KIMMEL, N. SOCHEN, AND J. WEICKERT, editors, *Scale-Space 2005*, **3459** of *LNCS*, pages 419–430. Springer Verlag, 2005. 28, 29, 36, 46

[139] ANUJ SRIVASTAVA, SHANTANU H. JOSHI, WASHINGTON MIO, AND XIUWEN LIU. **Statistical Shape Analysis: Clustering, Learning, and Testing**. *IEEE Transaction Pattern Analysis and Machine Intelligence*, **27**(4):590–602, 2005. 67

[140] J. STARCK AND A. HILTON. **Surface Capture for Performance Based Animation**. *IEEE Computer Graphics and Applications*, 2007. 97, 109

[141] GILBERT STRANG. **Maximal flow through a domain**. *Mathematical Programming*, **26**:123–143, 1983. 10

[142] C. STRECHA, R. FRANSENS, AND L. V. GOOL. **A Probabilistic Approach to Large Displacement Optical Flow and Occlusion Detection**. In *Statistical Methods in Video Processing*, LNCS 3247, pages 71–82. Springer Verlag, 2004. 44

[143] G. SUNDARAMOORTHI AND A. J. YEZZI. **Global Regularizing Flows With Topology Preservation for Active Contours and Polygons**. *IEEE Trans. Image Processing*, **16**(3):803–812, March 2007. 4

[144] S. R. THIRUVENKADAM, T. F. CHAN, AND B-W. HONG. **Segmentation Under Occlusions Using Selective Shape Prior**. In *Scale Space and Variational Methods in Computer Vision*, **4485** of *LNCS*, pages 191–202. Springer Verlag, 2007. 26, 44

[145] S. R. THIRUVENKADAM, T. F. CHAN, AND B. W. HONG. **Segmentation Under Occlusions Using Selective Shape Prior**. *Journal on Imaging Sciences*, **1**(1):115–142, 2008. 115

[146] A. TSAI, A. YEZZY, W. WELLS, C. TEMPANY, D. TUCKER, A. FAN, W. W. GRIMSON, AND A. WILLSKY. **A shape-based approach to the segmentation of medical imagery using level sets**. *IEEE Transactions on Medical Imaging*, **22**(2):137–154, 2003. 16, 26, 44, 60

[147] L. A. VESE AND T. F. CHAN. **A multiphase level set framework for image segmentation using the mumford and shah model**. *International Journal of Computer Vision*, **50**(3):271–293, 2002. xv, 8, 9

[148] B. VIJAYAKUMAR, DAVID J. KRIEGMAN, AND JEAN PONCE. **Invariant-Based Recognition of Complex Curved 3D Objects from Image Contours**. *Computer Vision and Image Understanding*, **72**(3):287–303, 1998. 69

[149] N. VU AND B. S. MANJUNATH. **Shape prior segmentation of multiple objects with graph cuts**. In *Computer Vision and Pattern Recognition*, pages 1–8, 2008. 115

[150] A.J. YEZZI AND S. SOATTO. **Stereoscopic segmentation**. In *IEEE International Conference on Computer Vision*, 2001. 103

[151] L. YOUNES. **Computable elastic distances between shapes**. *SIAM Journal of Applied Mathematics*, **58**:565–586, 1998. 69

[152] C. ZACH, D. GALLUP, J.-M. FRAHM, AND M. NIETHAMMER. **Fast Global Labeling for Real-Time Stereo Using Multiple Plane Sweeps**. In *Proceedings of Vision, Modeling, and Visualization*, 2008. 12, 13, 97, 98, 99

[153] A. ZAHARESCU, E. BOYER, AND R.P. HORAUD. **TransforMesh: a topology-adaptive mesh-based approach to surface evolution**. In *Proceedings Asian Conference on Computer Vision*, Tokyo, Japan, November 2007. 97, 109

[154] T. ZHANG AND D. FREEDMAN. **Tracking Objects Using Density Matching and Shape Priors**. In *International Conference on Computer Vision*, pages 1056–1062, 2003. 15, 16

[155] S. C. ZHU AND A. YUILLE. **Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation**. *IEEE Transaction Pattern Analysis and Machine Intelligence*, **18**(9):884–900, 1996. 3, 6, 13